Data Analytics for Automated Near Real-Time Detection of Blockages in Smart Wastewater Systems

Submitted by

Talia Rosin

to the University of Exeter as a thesis for the degree of

Doctor of Engineering in Water Engineering

in October 2020

Abstract

Blockage events account for a substantial portion of the reported failures in the wastewater network, causing flooding, loss of service, environmental pollution and significant clean-up costs. Increasing telemetry in Combined Sewer Overflows (CSOs) provides the opportunity for near real-time data-driven modelling of the sewer network.

The research work presented in this thesis describes the development and testing of a novel system, designed for the automatic detection of blockages and other unusual events in near real-time. The methodology utilises an Evolutionary Artificial Neural Network (EANN) model for short term CSO level predictions and Statistical Process Control (SPC) techniques to analyse unusual CSO level behaviour. The system is designed to mimic the work of a trained, experience human technician in determining if a blockage event has occurred. The detection system has been applied to real blockage events from a UK wastewater network. The results obtained illustrate that the methodology can identify different types of blockage events in a reliable and timely manner, and with a low number of false alarms.

In addition, a model has been developed for the prediction of water levels in a CSO chamber and the generation of alerts for upcoming spill events. The model consists of a bi-model committee evolutionary artificial neural network (CEANN), composed of two EANN models optimised for wet and dry weather, respectively. The models are combined using a non-linear weighted averaging approach to overcome bias arising from imbalanced data.

Both methodologies are designed to be generic and self-learning, thus they can be applied to any CSO location, without requiring input from a human operator. It is envisioned that the technology will allow utilities to respond proactively to developing blockages events, thus reducing potential harm to the sewer network and the surrounding environment.

Acknowledgments

Many people have contributed to the research that underpins this thesis. My sincere thanks to everyone who has helped me along the way.

First of all, huge thanks to my supervisors, Dr Michele Romano, Professor Zoran Kapelan and Professor Ed Keedwell, for their extensive support and technical input throughout the course of this project. I am grateful for their continual advice, patience, and attention to detail.

I would also like to thank United Utilities for their generous support in providing the data and the industrial knowledge, without which this work would not have been possible.

And finally, last but definitely not least, thank you to my family and friends for all their love, support and encouragement. My parents, brothers and sister who have lived with the ups and downs of this project for many years and did everything possible to see it to its conclusion.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

Sewer blockages are a major issue in the UK, and around the world. Blockages are responsible for the majority of service interruptions and flooding incidents which occur in the sewer network, and can cause numerous detrimental effects, including significant environmental pollution, damage to nearby properties and risks to public health. However, the increasing availability and accuracy of real-time sewer level data provides an opportunity for real time monitoring of the wastewater system.

This thesis describes the development of a novel blockage detection system, designed to detect blockages in the sewer system in real time. This chapter presents the motivation behind this thesis, first discussing the history of wastewater management practices and current blockage management practice. Next the aims and objectives of the work are outlined and the novel contributions it presents. Finally, an overview of the thesis's subsequent chapters is presented.

1.1 Wastewater Management: From Ancient to Modern Times

Waste treatment technologies have been devised to dispose of and treat wastewater since the dawn of human history. The issue has been dealt with in many different ways, and knowledge has been lost and regained over the centuries. During the Neolithic period (c. 10,000 B.C.E.) human communities were relatively small and nomadic, waste was returned to the land and decomposed using natural cycles. As permanent settlements developed other mechanisms became necessary to manage waste. This reached a peak with the Roman empire, who developed a largescale, complex sewer system which utilised wastewater for irrigation and fertilization purposes. There are also religious teachings from around this era that dealt with wastewater treatment. Mosaic Law (c. 1300 B.C.E.) instructs that one should "remove his own refuse and bury it in the earth", while the Talmud calls for the streets of Jerusalem to be washed daily.

However, many of the Roman advances in sanitation were forgotten during the Middle Ages (500 C.E. - 1500 C.E). In European cities waste was thrown into the

streets and open gutters for wastewater run-off ran along the centre roads. This led to rampant disease - in Europe it is estimated that at least 25% of the population died due to cholera, plague, and other water-borne diseases (Schladweiler, 2004).

The rapid growth of cities during the Industrial Revolution (1770s – 1840s) led to unbearable pollution caused by poor sanitation. During the early 19th century the River Thames was an open sewer. Outbreaks of cholera were endemic, resulting in a high death toll. The summer of 1858 saw the "Great Stink" in London, where hot weather exacerbated the smell of untreated human waste and industrial effluent polluting the River Thames. During this period English physician John Snow discovered the relationship between cholera and contaminated drinking water, tracing the disease to public wells polluted by wastewater.

This finding provoked a renewed interest into sanitation and promoted parliament to pass legislation enabling the creation of a modern sewerage system. Civil engineer Joseph Bazalgette was given responsibility for the work. His proposed design was an extraordinary feat of engineering, comprising of an extensive underground sewer system that diverted waste to the Thames Estuary, downstream of the main centre of population. The project was completed in 1875 and was capable of transporting 2 billion litres of waste every day. This system revolutionised the way that sewage is managed and brought forward a great

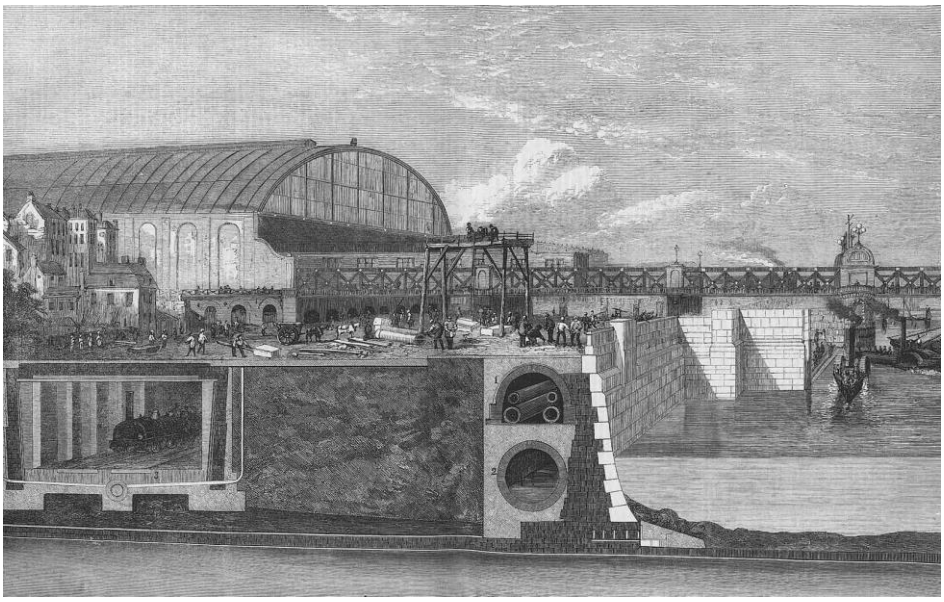


Figure 1-1 Cross-section of the Victorian Embankment, engraving, 1867

advance in sanitation. By the late 19th century sewer systems had been constructed in all the major cities in the UK. During the 20th century, major improvements were made to the sewerage system across the UK, including the construction of sewage treatment processes.

Bazalgette's sewer system is still in use today. Bazalgette had the foresight to design his network to accommodate a population increase of 50%. However, with substantial population growth, accelerated urbanisation and economic developments, the quantity of wastewater entering the system has increased significantly. The sewer system in the UK today is under severe strain. With climate change the proportion of rainfall occurring in high intensity events is rising, and sewers and treatment works are unable to cope with the large volumes of rainwater entering the system.

Combined sewer overflows (CSOs) are designed to protect downstream networks and wastewater treatment plants from hydraulic overloads and flooding during extreme rainfall events by discharging excess flow to nearby receiving watercourses (e.g. rivers, streams, estuaries, and coastal waters). However, these overflows consist of both foul sewage and stormwater and are a major source of pollution in urban waterways, harming the surrounding environment, degrading water quality and, threatening public health. The challenge now, therefore, is to create a resilient, sustainable and reliable wastewater network, whilst still maintaining affordable water bills.

1.2 Current Practice

Blockages in wastewater systems represent a major issue for water and sewerage companies in the UK today. There are approximately 300,000 blockages every year, resulting in costs of £100 million. These blockages account for 55% of sewer flooding incidents, causing over 3,000 flooded properties each year.

Blockages also increase the frequency and severity of CSO events. CSOs are designed to spill during heavy storms, when pollutants in the foul water are diluted, decreasing the impact on the environment. However, blockages and siltation can cause CSOs to spill at flow rates lower than consented. These

overflows, especially those that occur during dry weather flow, contain undiluted, heavily polluted untreated foul water. A large number of studies have been conducted analysing the various pollutants present in CSOs, and describing their significant detrimental impacts on the surrounding ecosystems and on drinking water qualities (Brokamp et al., 2017; Jalliffier-Verne et al., 2016).

Ofwat, the Water Services Regulation Authority, imposes a statuary duty on wastewater utilities in England and Wales to ensure that the functions of the sewerage system are properly carried out. However, with deteriorating sewer networks, an expanding population and increased water efficiency the number of sewer blockages suffered on public sewer networks is increasing. The timely detection and removal of blockage events therefore plays an important role in the management of urban water systems.

Historically, wastewater utilities have relied on their customers to report blockage events, employing a reactive repair and maintenance strategy. However, this approach means blockages are not identified until a visible event occurs – thus increasing the number of pollution events, the probability of loss of service as well as the number of customer complaints which occur. This in turn negatively affects the regulatory performance of the wastewater utility, which can lead to significant financial penalties. Utilities also employ hardware-based techniques to detect blockage events; CCTV is the current industry standard, and other method exists such as acoustic systems and laser profiling. However, these techniques are generally expensive, intrusive and time consuming to implement and require the added expensive of a human operator.

Efforts have been made to improve sewer system management using diverse modelling approaches. For example, physically based models are designed to model the hydraulic deterioration of sewer pipes based on an understanding of the physical mechanisms which control sewer pipe deterioration (e.g. Fenner et al., 2007). However, these physical mechanisms are generally not complex and not completely understood. As a result, the developed models are often too simplistic to reflect the actual failure processes and, furthermore, the data required to construct these types of models is often scarce. In addition, failures from blockages are not always dependent on pipe condition - deteriorated pipes

do contribute to blockage failures, however pipes that are in good serviceable condition are not immune to sever obstructions.

Another approach is to apply statistical models, which use historical data to predict future failures (Bailey et al., 2015; Xie et al., 2017). These models focus on understanding and predicting the risk of blockage events, structural defects and other sewer failures and represent their spatial distribution in the wastewater system using statistical and data mining techniques. These models are extremely useful as a risk analysis tool for directing sewer maintenance as sewer management resources, including funding, personnel, and access to assets, are often constrained. Statistical models can thus be used to predict which sections of the network are most likely to fail, enabling utilities to prioritise these areas for proactive management such as manual inspection, maintenance, and rehabilitation. However, these methods cannot detect the formation of actual blockages in real time. Therefore, there is still a need for reliable and accurate blockage detection technologies, capable of operating in real time, so that proactive maintenance can be quickly implemented before a blockage causes flooding or other adverse incidents.

In the last decade new sewer system sensors have becoming increasingly available and affordable, as have related information and communication technologies enabling near real-time wireless transmission of field data to company control rooms. In addition, the Event Duration Monitoring program, implemented by the Environment Agency, has requires wastewater utilities in England and Wales since 2020 to monitor sewer levels at the majority of their CSOs and report on their performance in terms of the number of discharges. This has resulted in the installation of large quantities of increasingly accurate level sensors throughout the sewer network. Driven by these regulations, and also due to reductions in data storage costs and improved computer processing power, wastewater utilities are now routinely collecting large volumes of accurate sewer level data in near real time. The amount of rainfall data available has also increased significantly, with the installation of the Met Office’s network of RADAR rainfall measurement network across the UK, which continuously provides rainfall readings at a resolution of up to 1km.

These substantial data sources provide an exciting opportunity for real-time data-driven modelling of the wastewater system. Data-driven methods allow patterns within the data to be derived, without the need for detailed hydraulic models of the sewer network. In addition to not suffering the disadvantages of physical models outlined above, these data driven models are also generally comparatively low cost, fast, and computationally inexpensive.

Real time data-driven models have been demonstrated to be effective and reliable when applied to water and wastewater modelling and have the potential to optimise the efficiency and to extend the functionality of the current sewer network. Real time control (RTC) models, for example, have been a research topic of great interest in the past decade. RTC systems dynamically control the flow and retention volumes in sewers using sensor networks and automated valves. They are viewed as a promising approach to reducing flooding and pollution incidents during rainfall events without the need for expensive investment in infrastructure. Several RTC systems have been described in the literature and have been successfully implemented in wastewater networks. Similarly, in the area of water distribution networks, many data-driven models designed for leak detection and localisation have been presented. These models make use of real time water pressure or flow data.

These data-driven approaches improve the management and operation of water and wastewater networks and have a beneficial effect on the environment, human health, and water conservation. However, few data-driven methodologies with the aim of blockage detection utilising real time sewer level data are available in the literature. The creation of an efficient and reliable blockage detection system will enable utilities to move towards a more proactive approach towards blockages events – i.e. removing obstructions before they cause damage to the sewer network and the environment and minimising the impact on the customer. The technology will also reduce operational costs significantly, by avoiding reactive ‘clean-up’ costs such as compensation, regulatory fines and penalties. The end result will be a more resilient and sustainable wastewater system.

1.3 Research Questions and Aims

1.3.1 Research Questions

The work in this thesis addresses the following key research questions:

1. Can a neural network model be designed to predict levels in a CSO chamber and to provide an early warning of spill events? If so, can the model be used effectively by wastewater utilities to better manage potential overflows?

2. Can blockage events be automatically detected from CSO level data, using a combination of statistical analysis and machine learning techniques? If so, can this be achieved in a reliable and timely manner, which can be applied across all types of sewer catchments and CSO chambers?

1.3.2 Research Aims and Objectives

The overall aim of the work presented in this thesis is to develop and evaluate a novel methodology to detect blockage which occur in the wastewater network. The system is designed to automatically and continually monitor the wastewater system for blockage events in real time using artificial intelligence and statistical analysis techniques. The detection system incorporates a model designed to forecast water levels in a CSO chamber and provide alerts for upcoming overflows. The detection methodology utilises only industry standard CSO level data and radar rainfall data. Thus, the technology can be easily implemented by any wastewater utility, without the added costs and inconvenience caused by the installation of additional hardware.

The specific objectives are as follows:

1. To perform a literature review of blockage detection in the wastewater sector and identify key gaps in knowledge.

2. To develop an AI model designed to forecast level in a CSO chamber and provide alerts for upcoming spill events. The methodology should utilise only standard CSO level data collected by monitors installed in CSO

chambers and rainfall data as a data source. The model should be able to generalise well to all CSO environments.

3. To determine the feasibility of detecting blockages and other unusual events which occur in the sewer system from CSO level data.

4. To develop a method for the automated detection of blockages and other unusual events in the sewer network in the proximity of CSO chambers. Again, the methodology should utilise only CSO level data and rainfall data. The proposed methodology should be suited to the needs of a field technician and be easily integrated into existing technology.

5. To test, evaluate, and demonstrate the performance of the CSO level prediction model at forecasting CSO levels and detecting upcoming spill events.

6. To test, evaluate, and demonstrate the performance of the blockage detection methodology on real blockage events in the UK. This should reflect real practice and evaluate the accuracy and efficiency of the detection system.

1.4 Thesis Contributions

The main contributions of this thesis are:

1. The development of a methodology for pre-processing of historic and real time CSO level data and rainfall data. This involved the development of a methodology to identify and remove benching from CSO level data.

2. The development of a bi-model committee evolutionary artificial neural network (CEANN) methodology for the prediction of water level in a CSO chamber and the generation of alerts for upcoming spill events. The model is composed of two evolutionary artificial neural network (EANN) models, optimised for wet and dry weather respectively, and combined using a non-linear weighted averaging approach to overcome bias arising from imbalanced data.

3. The development of a statistical analysis-based methodology to detect blockage events in near real time based on identifying deviations from normal CSO level behaviour.

4. The development of an EANN discrepancy-based methodology to detect blockage events in near real time based on the deviations between the EANN model predictions and observed CSO level data.

5. The integration of the above methodologies into a single event detection system. This involved the development of an inference engine designed to combine the event detection evidence and determine if a blockage has occurred. The system was applied to several real blockage events from a UK wastewater network.

1.5 Thesis Structure

The remainder of this thesis is structured as follows:

Chapter 2: Literature Review

This chapter presents a review of the relevant literature. The review gives an overview of the current techniques employed by wastewater utilities in the field of sewer system modelling and blockage detection.

Chapter 3: CSO Level Prediction Model Methodology

The first methodology chapter introduces the CSO Level Prediction Model. This model is composed of a CEANN. The model is designed to predict water level in a CSO chamber up to 6 hours ahead and provide alerts for upcoming spill events. Techniques designed to overcome data-imbalance are described. Three additional comparison ANN model are presented, which were developed in order to evaluate the performance of the CEANN model. The chapter also describes a data processing methodology designed to effectively process incoming CSO level data.

Chapter 4: Blockage Detection Methodology

This chapter presents the novel automated blockage detection system methodology. The system utilises AI and statistical analysis techniques to identify

anomalies in real time sewer level data. The chapter also describes the results of several data analyses used to design the system methodology and to select the detection system parameters.

Chapter 5: CSO level Prediction Case Study

In Chapter 5 the performance of the CSO level prediction model is demonstrated and evaluated on four case study CSO sites. This is done using historical CSO level data. Comparisons are made with the performance of the alternate ANN models. The chapter also presents several analyses conducted to demonstrate the capability of the model to provide warnings for upcoming overflow events.

Chapter 6: Blockage Detection Case Study

This chapter demonstrates and evaluates the capabilities of the blockage detection methodology when applied to real blockage events from a UK wastewater network.

Chapter 7: Summary and Conclusions

The final chapter summarises the material developed and discussed throughout this thesis and presents the relevant conclusions. The novel aspects of the work are outlined. Finally, potential directions for future research to extend and enhance the methodology are presented.

1.6 Publications Arising from Thesis

Rosin, T., Romano, M., Woodward, K., Keedwell, E., & Kapelan, Z. (2017). Prediction of CSO chamber water levels using rainfall forecasts*. Computing and Control for the Water Industry.*

Rosin, T., Romano, M., Kapelan, Z., Keedwell, E., & Woodward, K. (2018). Prediction of CSO chamber level using Evolutionary Artificial Neural Networks. *13th International Conference on Hydroinformatics.*

Rosin, T., Romano, M., Keedwell, E., & Kapelan, Z. (2019). Near Real-Time Detection of Blockages in Wastewater Systems using Evolutionary Artificial Neural Networks and Statistical Process Control. *10th International Conference NOVATECH.*

Rosin, T., Romano, M., Keedwell, E., & Kapelan, Z. (2019). Data Analytics for Automated Detection of Blockages in Sewers. *17th International Computing and Control in the Water Industry Conference.*

Rosin, T. R., Romano, M., Keedwell, E., & Kapelan, Z. (2021). A Committee Evolutionary Neural Network for the Prediction of Combined Sewer Overflows. *Water Resources Management*, 35, 1273–1289. DOI: 10.1007/s11269-021-02780-z.

Rosin, T. R., Romano, M., Keedwell, E., & Kapelan, Z. (Submitted). Near Real-Time Blockage Detection in Wastewater Systems Using Data Analytics.

Chapter 2: Literature Review

2.1 Introduction

This chapter provides a critical review of the literature relating to sewer system modelling and blockage detection. The aim is to investigate the previous work carried out in the field and identify gaps in knowledge that will be addressed in this thesis. The modelling of blockages is relatively limited in the literature, especially regarding real-time data-driven modelling. Therefore, the review also describes methods from related fields, such as fault detection in water distribution networks, which have the potential to be applied to wastewater blockage detection.

The chapter is organised as follows: first an overview of blockage events in sewer systems is presented in Section 2.2. This aims to provide a picture of blockages and their formation, describe current management by wastewater utilities, and highlight the need for reliable blockage detection methods. Next an overview of combined sewer overflows (CSOs) is given in Section 2.3, describing the role they play in the combined sewer systems and their relation to sewer blockages. A discussion on the increasing availability of sewer network data and rainfall data is then presented in Section 2.4. The review then moves on to CSO modelling and sewer blockage detection techniques. Section 2.5 presents an overview of CSO modelling, describing physical and data-driven models. Section 2.6 then describes blockage detection techniques, covering hardware and software techniques, outlining the capabilities and limitations of each technology and evaluating their potential, as well as giving a brief overview of fault detection and diagnosis in engineering systems. Relevant fault detection and modelling methods found more generally in the water sector are then presented in Section 2.7. Section 2.8 discusses blockage detection systems available commercially and presents a brief overview of various projects implemented by wastewater utilities. Finally a discussion and summary of the chapter, containing the main conclusions and considerations, and highlighting the gaps in the current research is given in Section 2.9.

2.2 Sewer Blockages

2.2.1 Blockages Overview

Blockages are responsible for the majority of incidents in wastewater networks, as well as a large proportion of the cost. The sewer system in the UK is approximately 300,000 km long. According to Ofwat data for 2008/2009 over 200,000 sewer blockages were reported in the UK during this time period. This indicates an average blockage rate of 517 blockages per 1000 kilometres of sewer per year (Hillas, 2014).

Sewer blockages have a significantly detrimental effect on the urban drainage system. They cause problems such as external and internal flooding, odour, pollution of natural watercourses and contamination of gardens and roads. Indeed, it is estimated that blockages are the main cause of sewer serviceability loss in both dry and wet weather flow conditions. During the period of 2008/2009 approximately 2% of blockages resulted in internal flooding of properties and 23% in external flooding. These events cause distress and inconvenience to residents and pose potential health risk due to waterborne pathogens (Veldhuis et al., 2010). Flooding is very costly to the water industry, incurring regulatory fines, maintenance work, and clean-up costs.

The Water Services Regulation Authority (Ofwat) is responsible for economic regulation of the privatised water and sewerage industry in England and Wales. They impose a statutory duty on wastewater utilities to maintain the condition and serviceability of their assets, utilising incentives and penalties to encourage companies to improve their service in terms of quality and value to customers. Ofwat's PR19 price review (Ofwat, 2017b), which describes the goals and commitments water and wastewater companies will be measured under for 2020-2025, emphasises the need for resilience of the water and wastewater networks. Of the 14 common performance commitments which all wastewater companies are required to adhere to, three are directly related to the improved management of sewer blockages: reduction of sewer flooding, reduction of sewer collapses and reduction of pollution incidents (Ofwat, 2017c). Additionally, as part of the asset health performance commitment, wastewater utilities can be required to record the number of blockages which occur in their networks and cause a

reportable problem (Ofwat, 2017a) and failing to meet these commitments can incur severe penalties. Severn Trent Water for example, who are responsible for water and sewerage services in the Midlands of England, will be fined £11200 per blockage event if they underperform on their sewer blockage performance during the PR19 period, and will be rewarded £3700 per blockage if they overperform (Ofwat, 2019).

These regulations incentivise wastewater utilities to reduce the number of blockages which occur in their networks and to minimise their impacts. However, blockage management is a difficult task. The wastewater network is a complicated system, containing a large number of sewer pipes, with generally low telemetry coverage and a small rate of blockages compared to the size of the system. Sewer maintenance must accommodate many different engineering, social, environmental, and economic constraints. As sewer pipes are underground, sewer maintenance operations are generally costly and disruptive. Additionally, the urban drainage system in the UK suffers from a combination of aging infrastructure, insufficient capacity, a lack of maintenance, an increasing population, and increasing urbanisation. Water conservation, which is required due to water stress, results in less water in the system, in turn reducing solids transfer in sewers. However, climate change is also causing more extreme weather patterns, putting additional pressure on the sewer network during intense storm events (Salerno et al., 2018).

Traditionally, water utilities have approached blockage maintenance (and sewer maintenance in general) reactively, relying on customers to call up and report a complaint. This can be for example, a report of internal flooding, a restricted toilet, or the presence of odour. However, the resulting cost of sewer failures caused by these blockages, including service disruptions, bad publicity, and health and safety problems, is significant. Additionally, this approach can lead to a high number of customer complaints.

The current trend in urban water systems management, therefore, is moving towards a more proactive approach, i.e., trying to address potential problems before they occur. A number of studies have shown that implementing a proactive maintenance program, such as regular sewer cleanings, is a cost effective method of reducing the frequency of service disruptions and their undesirable

consequences (flooding, pollution, etc), and minimising potential environmental damage (e.g. Fontecha et al. 2016, Ashley et al. 2000 and Veldhuis & Clemens 2011). Additionally, preventive maintenance programs can extend the lifetime of equipment and infrastructure, an important issue for utilities as replacement and reparation of the sewer system is expensive and disruptive.

A study by UKWIR in 2010 found that approximately 75% of the blockage management activity was in reaction to failures (UKWIR, 2010). However, it is likely the proportion of proactive maintenance has grown in the last decade. This increase is enabled by the advancement of new low-cost sensors, the growing use of the Internet of Things (IoT), and the potential of social media platforms to report different types of problems.

2.2.2 Blockage Formation

Blockages occur in sewer networks due to a build-up of material in the sewer which restricts sewer flow. Blockages can be caused by a large variety of factors; the most common mechanisms are:

- Failure of the gross solids transport mechanism in the sewers, e.g. due to discharge of unsuitable solids such as rags, nappies and sanitary products, insufficient gradient, or a defect in the pipe
- Blockages at interceptor traps in drains or sewers
- An accumulation of solid fats, oils, and grease (FOG)
- Obstruction of the pipe by tree root intrusion through joints or defects in the sewer wall
- Other foreign obstructions (e.g. bricks, concrete, other utility services etc).
- Sedimentation

Understanding the composition of blockages can aid in managing and reducing blockage events. Prior to the 1990's there were relatively few studies on wastewater blockage events. A report by the Institution of Sanitary Engineers (Institution of Public Health Engineers, 1954) identified the immediate causes of blockages as; sanitary towels (37%), newspapers (23%), rags (11%) and grease (5%). It also found that a large proportion of blockages were associated with interceptor traps. Lillywhite & Webster (1979) analysed blockages in 100 mm

and 150 mm diameter pipes, identifying the main blockage causes as defective pipe joints (30%) and deposits in the line (23%). However, it is important to note that the study only analysed 70 blockage events.

Marlow et al. (2011), analysed blockage data from 2 Australian water companies over a period of 4 and 9 years and found that the majority of blockages were caused by tree roots (including blockages where FOG has accumulated in the tree roots), making up 67% and 72% of all blockage events for the two different water companies. Blockages were also caused by damaged pipes (3% and 0.6%), FOG (21% and 8%) and unknown causes (9% and 12%). These findings are similar to a survey conducted in Poland by Kuliczkowska (2008) which found that tree roots were responsible for between 50% and 78% of all sewer blockages, depending on the material of the sewer pipe.

Recent reports from many water companies have also indicated that Non-Flushable Products (NFPs) such as sanitary towels and wipes have become a major cause of blockage events. A Water UK report (Water UK, 2017) found that over 75% by weight of identifiable products recovered from blockage material were baby wipes. Surface wipes, cosmetic removal wipes and feminine hygiene products accounted for a further 20%. Products designed to be flushed accounted for a very small proportion - approximately 0.88% by total weight and 1.9% by weight of products that could be identified. From this they inferred that a significant number of people are either unaware of the ‘do not flush’ advice, do not appreciate the reason why wipes should not be flushed, or are unconcerned by the potential consequences.

Blockages due to FOG have also become a significant global problem in urban areas. FOG enters the sewer system from a variety of sources including households, food service establishments and food processing factories. Due to increasing populations and changing diets, the amount of FOG entering the sewer system has increased over the past decade. In some parts of the UK FOG is reported to be responsible for 75% of blockage events (Martin, 2017), resulting in annual control costs between £15 million and £50 million (Del Mundo & Sutheerawattananonda, 2017). FOG attracts insects and vermin such as rats, and sloughed deposits can affect pumping stations and wastewater treatment plant operations.

FOG also compounds other blockage causes, as it reacts with materials such as wet wipes and nappies in the sewer, creating an insoluble solid material which obstructs the pipe. This can result in the formation of so called ‘fatbergs’ – giant FOG blockages. This term was initially coined by media when the first fatberg was found in Kingston upon Thames in 2013. In recent years the reporting of fatbergs has been prevalent - numerous global media outlets such as the BBC, the New York Times and National Geographic have produced articles on the topic (BBC News, 2017; National Geographic, 2017; The New York Times, 2019).

The largest fatberg in the UK was removed from the sewer network in Whitechapel, East London, weighing 130 tons and measuring 240 meters in length. The removal costs and repair of the sewer network was estimated to be around £400,000 (Moore-Bridger 2017; Adams 2018). Many cities have instigated educational campaigns and inspections of food service establishments in an attempt to reduce the amount of FOG entering the sewer system. However FOG blockages remain a significant concern.

The final blockage cause, sediment, consists of fine organic or inorganic particles which accumulate in the sewer pipe. Sediment is found in all types of sewer system, although the actual makeup of the sediment depends on whether the sewer system is foul, combined or surface water. Crabtree (1989) describes five classes of sewer sediment based on their size, chemical composition, and deposition patterns. Effective sewer systems are designed according to self-cleaning criteria, where the sediment is transferred within the sewer flow to the wastewater treatment works. However, if the velocity in the sewer is too low, either permanently or intermittently, the sediment is deposited by the flow and accumulates on the bottom of the sewer pipes. This generally occurs in sewers with a low gradient or a high diameter to flow ratio, or sewers which are in bad condition. Sedimentation restricts sewer flow and therefore must be flushed periodically.

2.3 Combined Sewer Overflows

Combined sewer systems carry both wastewater and stormwater in one unified sewer network to the wastewater treatment plant (WWTP). This is in contrast to separate systems where the wastewater and storm water are transported in two

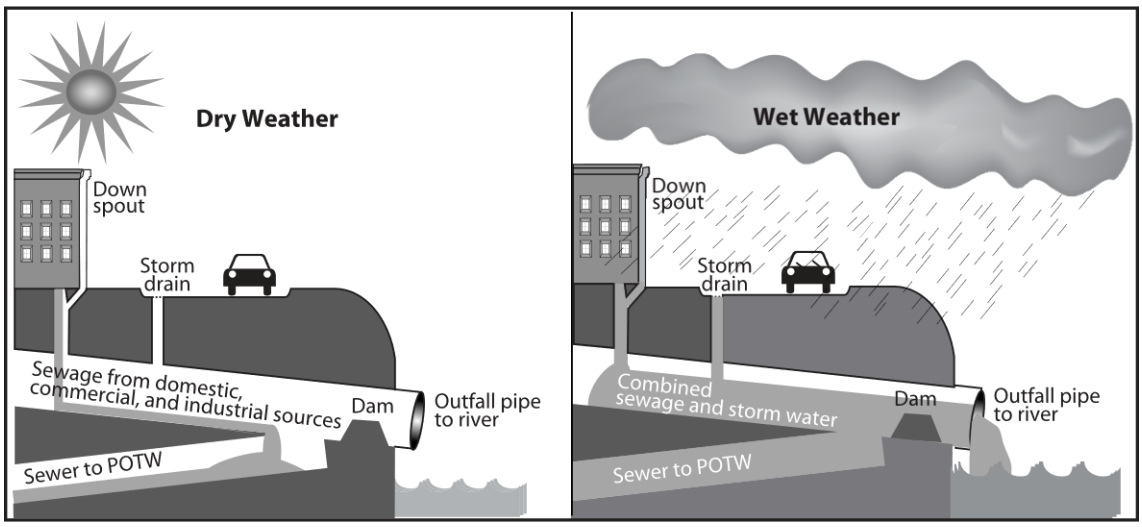


Figure 2-1 Illustration of a combined sewer overflow (from Wikipedia, n.d.)

separate pipelines. Combined systems make up a significant proportion of sewer networks in many cities around the world. In central Europe, approximately 70%of the sewer system is combined (Butler et al., 2018), resulting in approximately 650,000 CSOs. in the USA a total of 9348 CSO outfalls have been identified, distributed across 32 states and nine Environmental Protection Agency Regions (Botturi et al., 2020).

CSOs are a necessary component of combined systems, designed to protect downstream networks and wastewater treatment plants from hydraulic overloads and flooding during extreme rainfall events. An example of a CSO operating during dry and wet weather is illustrated in Figure 2-1. They operate by diverting excess flow beyond the capacity of the network to a nearby receiving watercourse (e.g. rivers, streams, estuaries, and coastal waters). However, as these discharges contain both foul sewage and stormwater they contain large amounts of pollutants in the form of gross and finely suspended solids and pollutants in solution, as well as household items such as wet wipes, sanitary items and cotton buds. As a result, the overflows are a major source of pollution in urban waterways and can cause significant harm to surrounding environments and ecosystems, degrading water quality, threatening public health and resulting in regulatory failures.

CSOs are designed to spill only during heavy storm events when pollutants in the foul water are diluted, thus decreasing the impact on the environment. However, incidents such as blockages and siltation can cause CSOs to spill at flow rates

lower than consented. These overflows, especially those that occur during dry weather flow, contain undiluted, heavily polluted untreated foul water. A large number of studies have been conducted investigating the various pollutants present in CSO spills, and analysing their harmful impacts on the surrounding ecosystem, UK marine life and drinking water quality (García et al., 2017; Jalliffier-Verne et al., 2016; Wei et al., 2019). According to Phillips et al. (2012) micropollutant concentrations could be up to 10 times higher in CSO discharges than in treated wastewater. Recent studies indicate that CSO spills contribute from 30–95% of the annual load for a number of different pollutants, while contributing only a small proportion of the total annual wastewater discharge (Launay et al., 2016, Phillips et al., 2012).

The European Environment Agency's 2019 overview of bathing water quality across the EU found that only 66.1% of the UK's bathing waters were of excellent quality, compared with an EU average of 84.8% (EEA, 2020). As stated above, climate change predictions indicate an increase in the frequency and intensity of extreme rainfall events in the near future – indicating that CSO events will occur even more frequently.

Wastewater utilities are under increasing pressure to reduce CSO events. They are required to abide by a number of regulations regarding CSOs, including the EU Water Framework Directive (Council Directive (EC) 2000/60/EC) and the European Urban Wastewater Treatment Directive (EU UWWTD, 1991). Ofwat guidelines (Drainage Strategy Framework, 2013) for recommended good practice state that water utilities should:

- Improve understanding of network performance (and improve models) by using long term flow and level monitoring on sewers and CSOs.
- Establish systems and apply methods to predict how interventions can reduce risks, e.g. how reducing infiltration can reduce CSO spills.

If CSO events can be predicted in advance there are techniques for mitigation, such as maximising storage, and adjusting set points for movable gates and pumping stations. Consequently, there has been increasing interest in monitoring CSOs to provide an early warning for overflows in the past decade. A variety of techniques have been developed for predicting and managing spill events, which

are described in more detail in Section 2.5. In addition, managing the large number of sewer blockages has a major role to play in reducing CSO volumes and improving the water quality standards of rivers and coastal waters.

2.4 Data Availability

Recent years have seen large developments in hydraulic sensor technology. Level, flow, and pressure sensors have become increasing accurate and inexpensive. These sensors collect data at low sampling frequencies - standard industry practice in the UK is to sample at regular intervals of 15 minutes. Advances in on-line data acquisition and telemetry systems, as well as decreasing communication costs, have also enabled utilities to collect long term asset performance data.

In particular, the monitoring of CSO chambers has increased significantly. This has been driven by the introduction of the Event Duration Monitoring (EDM) program, announced by the Environment Agency in 2014. The program requires monitoring at the majority of CSOs in England and Wales from 2020, with the aim of improving the visibility of the performance of sewer networks. As flow rate and total overflow volume are relatively difficult and expensive to monitor, for the majority of CSO sites this takes the form of level monitors used to log the timings and duration of spill events (CIWEM, 2016). Measurements are typically made at 2 or 15-minute intervals. In a minority of cases, flow sensors may be installed to check the compliance of individual CSOs against consents set by the Environment Agency (EA). The Chartered Institution of Water and Environmental Management provides a good practice guide (CIWEM, 2016) on the practical application of CSO monitoring. As a result, water and wastewater utilities in the UK have begun installing large quantities of level sensors in their networks, routinely collecting large volumes of sewer level data. The data is often collected at the rate it is measured, i.e. in near real time.

There have also been significant advancements in the study of rainfall radar data. In the past, rainfall data was limited to rain gauges. Range gauges accurately measure rainfall on the ground at point locations, however there may be many kilometres between each gauge and data must then be interpolated between measuring stations. Studies have shown that the information obtained from one

gauge does not always provide a good representation of rainfall amounts in the whole area (Moore, 2014). In recent years rainfall radar data, collected by a network of C-band radars in the UK, has become widely available. Radar rainfall data enables detailed detection of spatial and temporal rainfall patterns. Combining radar data and rain gauge data is a considerable improvement to using rain gauge data alone. Liguori et al. (2012) demonstrated that radar data provides valuable rainfall measurements which can be used in hydraulic modelling. Since then many studies have used rainfall radar data, including hydraulic models and machine learning approaches for predicting urban flooding in real time, e.g. Duncan et al., 2013, Mounce et al., 2014a.

Radar rainfall data for the UK is processed by the Met Office, who currently provide data with a resolution of at least 2 km for over 85% of the country (which includes most of the large urban catchments) (Met Office, 2009). Rainfall radar data are generally supplied in near real-time at resolutions of 5 min.

These large volumes of detailed level and rainfall radar data provide opportunities for in-depth analysis of the wastewater system, and the potential for monitoring of the network in real-time. In particular it supports the development of data-driven modelling techniques, which allow patterns within the data to be used to derive relationships, without the need for the construction of detailed hydrodynamic models. Examples of these types of models are detailed in the following section.

2.5 Combined Sewer Overflow Modelling and Management

As stated in Section 2.3, minimising CSO volumes in urban areas is an important task. In many countries the existing sewer networks are not designed to handle the combined stormwater and wastewater volumes which occur during heavy rainfall events. Moreover, climate change is causing increasingly intense precipitation events in certain areas, resulting in more frequent spills. These events can cause significant environmental risks if they are not properly controlled. Implementing methods to better manage CSOs is therefore an important task. In recent years there has been an increase in the literature covering CSO management and describing methods designed to reduce the impacts of CSO spills on receiving waters, in terms of both quantity and quality.

CSO volumes can be minimised using either structural or non-structural methods. Structural methods refer to physical constructions built into the sewer system and designed to reduce overflows. These include underground tunnels and storage tanks used to store combined sewer flows during heavy rainfall. Non-structural methods refer to approaches designed to reduce overflow volumes in existing sewer networks without implementing any physical structures, for example control algorithms based on optimization theories. Structural methods are affective and can be implemented when the space is available. However they are generally expensive and disruptive to the surrounding environment (Zhang et al., 2018a). Consequently, research and investment into non-physical methods is often preferred.

However, designing non-structural solutions has many challenges. The wastewater network is a complex system due to the dynamic behaviour of sewer flow, and the hydrological response is often fast and sensitive to precipitation variability at small scales. The accuracy of CSO models is strongly dependent on current and future rainfall data, which is generally only available with adequate precision for short lead times. The best methods for modelling CSOs in real time thus provide updates every few minutes using the most recent rainfall nowcasts and data obtained through network monitoring.

The following section describes the various methodologies designed for modelling and management of CSOs, covering both physical and data-driven techniques.

2.5.1 Physical Modelling

Wastewater utilities have traditionally constructed physical models of the sewer system. However, physical modelling of water levels in the sewer system is a difficult task. Models must take into account multiple interactions in the various sub-systems including catchments, sewer pipes, wastewater treatment plants and receiving water bodies (Saagi et al., 2016, 2018). Physical models require detailed information of the sewer system and incorporate many parameters which are not always available, they are often difficult to build and calibrate, and are computationally expensive.

The preferred physical model for water motion in sewer network simulations is based on the Saint-Venant equations, a set of hyperbolic nonlinear partial differential equations, which relate flow and water level in an open channel (Marinaki & Papageorgiou, 1998). The Saint-Venant equations are used to construct calibrated hydraulic models which are fed with real or design rainfall. However, the equations are generally not suitable for use in real time as they contain a high number of variables and are nonlinear in nature and so have high computational costs. Therefore, simplified mathematical models have been developed which are able to describe flow with reasonable accuracy and can be solved in a short amount of time.

Another popular approach is the conceptual virtual tank model, first developed by Gelormino & Ricker (1994) which is designed for optimisation based control of large-scale sewer networks. This technique has been applied by a number of studies (Joseph-Duran et al., 2014; Ocampo-Martínez et al., 2005; Puig et al., 2009; Svensen et al., 2019).

2.5.2 Data-Driven Modelling

In recent years data-driven modelling has become an increasingly popular and attractive option for modelling of the sewer system. Developments in computational intelligence, in the area of machine learning in particular, have greatly expanded the capabilities of data-driven techniques, allowing them to solve various complex problems. As the name suggests, data-driven methods analyse historical process data to produce knowledge about a system. No physical information concerning the underlying processes is required, making them well suited to situations where theoretical models of behaviour are poorly developed or inadequate, and where training data is plentiful or easy to collect. Data-driven models are also often much less computationally expensive than physical methods, and so are suitable for running in real-time. With the increase in the availability of real-time sewer level data, data-driven models are thus well suited to sewer network modelling, and hydraulic modelling in general. A comprehensive review of hydrological data-driven modelling is presented by (Remesan & Mathew, 2015).

Despite their good performance data-driven models do have a number of drawbacks, however. Data-driven models often require large amounts of training data to represent both normal and faulty modes of operation. The models are also often specific to the system for which they are trained and so cannot be applied to other systems.

Popular data-driven approaches for modelling and fault detection in the water and wastewater sector include wavelets, artificial neural networks, fuzzy logic, and support vector machines. Artificial neural networks and evolutionary artificial networks are described below in more detail as they are utilised prominently in the work presented in this thesis:

Artificial Neural Networks

Artificial Neural Networks (ANN) are a biologically inspired computational modelling approach based on the neural systems found in the human brain. The network architecture contains a set of interconnected processing elements called nodes, or neurons, which are trained to represent the relationships and processes inherent within system data. The neurons receive an input signal, pass it through an activation or transfer function such as a logistic or sigmoid curve, and transmit an output signal to other interconnected neurons. The neural network accumulates knowledge at each layer of the model through a self-learning process until it can capture and accurately reproduce the behaviour of the system and forecast future behaviour (House-Peters & Chang, 2011).

Many different types of neural network models have been developed in the last few decades; the most commonly applied structure is the feed forward hierarchical architecture trained using the back-propagation method. Feed forward indicates that the data flow is unidirectional from the input to the output units. The backpropagation algorithm is a variation of a gradient descent optimisation algorithm which minimises the error between the predicted and actual output values. The weighted connections between the neurons are adjusted after each training cycle until the error in the validation data set begins to increase. If this approach is not applied the network may represent the training data set too well and will be unable to generalise to unseen data (Solomatine et

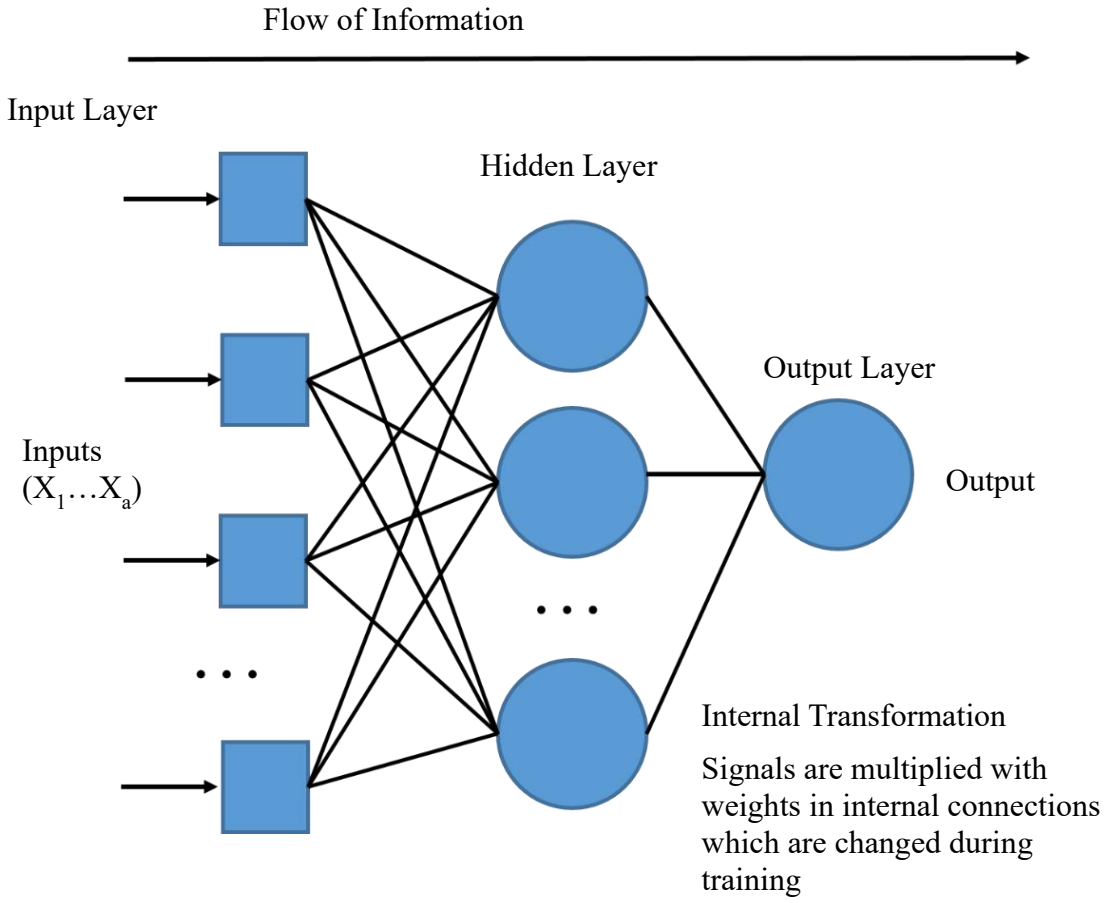


Figure 2-2 Neural network architecture.

al., 2008). An example of a feedforward single hidden layer ANN is illustrated in Figure 2-2. This topology is frequently used in hydrological applications.

The self-learning ability of ANN models allows them to approximate a wide class of non-linear functions and extract significant features or characteristics from complex data networks. Additionally, unlike many other methods, ANNs are not greatly affected by any error-corrupted values in the input data. However, ANNs work essentially as black boxes and attempt to develop a relationship among input and output variables without considering the physical processes involved. As a result, it is difficult to understand how neural networks arrive at their decisions and they cannot provide any insight into system behaviour. Additionally, the intelligence and expertise of human operators, which has often been built up over several years, cannot be easily incorporated into ANN models.

Over the last few decade ANNs have gained much attention and have emerged as a prominent tool for hydraulic modelling. A number of comprehensive reviews have evaluated the use of ANN's in hydrology. Evora & Coulibaly (2009) presented a review of advances in ANN modelling, focusing on remote sensing applications developed since 2000, and in particular those which use precipitation data. Li et al. (2010) reviewed the use of ANNs in urban hydraulics, including sewer flows, run-off sewer flows and CSO quality modelling. Maier & Dandy (2001) presented a review of 43 papers applying ANNs to the prediction and forecasting of water resource variables. The vast majority of these models utilise feed-forward networks and generally apply the backpropagation training algorithm.

Several ANN models have been demonstrated to successfully forecast CSO levels. Fernando et al. (2006) designed one of the earliest models; a feed-forward, back-propagation ANN used to forecast CSO events using flow rate and rainfall data. The model was designed to investigate the feasibility of replacing the existing traditional hydrological/hydraulic model serving Auckland City with an ANN type model. The ANN provided reasonably accurate sewer overflow rate forecasts. However, it's accuracy was highly dependent on the availability and quality of real-time flow data, which was also used to provide antecedent flow data. Without the antecedent flow rates the model forecasts were relatively poor.

Sumer et al. (2007) researched the viability of real-time detection of sanitary sewer overflows (SSOs) using time series analysis and ANN techniques. The methodology was applied to two case studies in Arizona, USA. An ANN was developed to estimate the 6-hour component of the forecast and to determine if an SSO was occurring control limit theory was used to detect important deviations between measured and expected depth and flow data.

In Kurth et al. (2008) a three hidden-layer feed-forward multilayer perceptron (MLP) was used to predict CSO levels 3 time steps (15 minutes) ahead using level and rain gauge rainfall data. The model utilised 12 past timesteps of rainfall data from 2 nearby rain gauges and 5 timesteps of past CSO level data. Guo & Saul, (2011) also used in-catchment rain gauge data, developing an adaptive linear ANN (named ADALINE) to model the relationship between CSO water level and rainfall.

Chiang et al., (2010) proposed a three-layer recurrent neural network with internal time-delay feedback loops to predict sewer levels in real-time. The ANN produced good outputs for 5 to 20 minute ahead predictions. The results were compared with the Storm Water Management Model (SWMM), a hydrological simulation model based on the Saint-Venant equations and commonly applied to designing quality and quantity models of runoffs in urban areas. The ANN model was found to produce favourable results.

More recent models have utilised radar rainfall intensity data rather than rain gauge data. The use of rainfall radar data reduces data collection requirements, which is particularly beneficial in large catchments where several rain gauges may be needed. Mounce et al. (2014b) investigated the use of radar rainfall intensity data when using an ANN to predict CSO depth 75 minutes ahead. They concluded that their model improved on previous studies using tipping bucket rain gauge measurements and demonstrated a prediction of CSO depth with less than 5% error for predictions more than one hour ahead for unseen data.

This model was further advanced by Mounce et al. (2014a) who removed the inputs of recent water depths for prediction, using only rainfall radar data. This approach resulted in increased errors, however, it allowed changes in the overall performance of the CSO to be better understood, which could then be used to identify abnormal CSO behaviour. Predicted and measured CSO levels were fed into a Fuzzy Inference System (FIS) developed which generated a traffic light coded assessment of CSO behaviour.

Zhang et al. (2018b) compared four different neural network models: a multilayer perceptron (MLP), a wavelet neural network (WNN), a long short-term memory (LSTM) and a gated recurrent unit (GRU), determining that the GRU model produced the best CSO level forecasts. Zhang et al. (2018a) used Multi-task Deep Learning to forecast CSO events from multiple CSO structures simultaneously at a citywide level.

More generally in wastewater drainage many ANN models have been successfully applied to areas such as sewer flow modelling (She and You 2019; Zhang et al. 2019), urban flooding (Berkhahn, Fuchs, and Neuweiler 2019; Kim and Han 2020), modelling of sediment transport and accumulation (Al-Ani & Al-Obaidi, 2019; Ebtehaj & Bonakdari, 2014) and water demand forecasting

methods (Adamowski, 2008; Bougadis et al., 2005; Ghiassi et al., 2008; Odan & Reis, 2012).

Evolutionary Artificial Neural Network

When utilising ANN models the selection of the model architecture and inputs is of great importance and can have a significant impact on the model accuracy. The process used to construct ANN models is typically empirical; the designer follows a practical standard or uses experimental insight in conjunction with trial and error (López-Vázquez et al., 2020). This trial-and-error approach generally results in the selection of effective variables; however it is a time-consuming, labour-intensive process and in addition requires a human expert.

Sewer network often contain very large numbers of CSOs. Therefore constructing individual ANN models via trial and error for each site is not realistic. However, different CSOs often exhibit very different behaviours and responses to rainfall in their catchments. Thus ANN models cannot just be transposed from one site to another and guaranteed to produce good results.

A solution to this problem is to utilise an evolutionary artificial neural network (EANN) model. EANNs refer to a class of ANN whereby evolutionary algorithms (EAs) are used in the ANN model design and/or training. The EAs are optimised to automatically determine the optimum parameters, encoding schemes, fitness functions, evolution processes, training, architectures, learning rules, etc. of the ANN (Yao, 1993). The use of EANNs allows a closer to 'optimal' results and also removes, partially or totally, the requirement of a human expert to construct the model.

EAs are a class of stochastic search and optimization techniques inspired by the process of natural selection in biological systems and are designed to perform searches over complex spaces (Goh & Tan, 2009). The term EA covers a broad spectrum of approaches, which make use of techniques such as inheritance, mutation, selection and crossover. Notable examples of EAs include genetic algorithms (GA) (Holland, 1984), evolutionary strategies (ES) (Schwefel, 1998; Tripathy & Schwefel, 1982), differential evolution (Storn & Price, 1997) and genetic programming (Koza, 1994).

The main principles of all types of EAs are the preferential survival and reproduction of the fittest members of the population. Given a quality function to be maximised, a random set of candidate solutions is created. By applying this quality function as an abstract fitness measure only the fitter candidates are chosen to survive and proliferate, while the unfit candidates die off and do not contribute to the gene pool of further generations. Thus resulting in a rise in the fitness of the overall population. Mutations can be applied to the candidates, increasing the genetic diversity in the candidate pool. This process is iterated until a candidate of sufficient quality is identified or a computational limit is reached.

EANNs have been used successfully since the 1990s. Yao (1993, 1999) produced two comprehensive reviews, examining the different combinations between ANNs and EAs, exploring the use of EAs when evolving connection weights, architectures, learning rules and input features, and reviewing the different search operators used in various EAs. There has been increasing interest in applying EANNs to hydraulic systems in the last decade. For example Chen & Chang (2009) demonstrated that an EANN can be used to forecast 10 day reservoir flow, with an accuracy greater than autoregressive models. Chaves & Kojiri (2007) applied a conceptual fuzzy neural network (CFNN) to a water quality model in a reservoir system, using a genetic algorithm (GA) to find the parameters of fuzzy interference and the connection weights. Chaves & Chang (2008) developed the Evolving ANN Intelligent System (ENNIS), an intelligent system designed to optimise the operation of a multi-purpose reservoir. Moradi & Dariane (2017) proposed an evolving neural network to derive operating policies for complex reservoir operations. However, to our knowledge, EANNs have not yet been applied to CSO level forecasting.

2.5.3 Real Time Control Systems

Real time control (RTC) refers to active control of the wastewater network by operating actuators such as movable control valves, weirs, gates and pumps to make more effective use of existing storage capabilities in the sewer network and improve flow regulation. This enables the system to adapt to changing rainfall conditions and sewer behaviour in real-time. RTC is a cost-effective strategy as it improves sewer performance based on the existing draining facilities, helping to prevent the need for the construction of additional costly retention tanks in the

sewer network. This is especially beneficial in densely populated areas with limited space for new constructions. RTC can be viewed as part of the emerging concept of ‘smart cities’, whereby cities develop from being static to flexible systems, utilising data which is monitored, analysed and managed in real time (Lund et al., 2018).

Research into RTC is not a new phenomenon; it has been a topic of interest in the field of urban wastewater systems since the 1960s (Borsányi et al., 2008). However, with the recent increased installation of monitors in the sewer system, RTC of wastewater systems has seen a significant increase in studies in the last decade. RTC systems most commonly use water level sensors, due to their affordability, durability and accuracy. However, less frequently flow meters are utilised.

A number of RTC systems have been developed with the aim of reducing flooding and combined sewer overflows, utilising both physical (Altobelli et al., 2020; Garofalo et al., 2017) and data driven techniques (Gu et al., 2001, Regneri et al., 2013, J. Li, 2020, Mounce et al, 2020).

2.6 Blockage Detection Techniques

Blockage detection methods can be categorised into two classes - hardware and software-based techniques. Hardware-based methods utilise sensing devices and other equipment to detect blockages and other unusual events, while software-based methods use real-time monitoring data, and include hydraulic model-based methods, transient-based methods, and data mining methods. The following section gives a brief overview of the most commonly applied hardware and software-based methods.

2.6.1 Hardware Based Blockage Detection Techniques

Hardware based techniques are a popular and commonly used method for inspecting sewers and identifying the presence of blockages in wastewater pipes. There are several hardware techniques available for the detection and localisation of blockages. As sewer infrastructure grows there is increased

pressure for companies to conduct more extensive and detailed inspections of their underground pipe systems.

However, sewer systems have many characteristics which make inspection difficult; for example the lack of light inside the sewer pipes can negatively affect visual sensor measurements, water levels inside the sewer pipe vary significantly from dry to completely flooded, and sewer walls are generally rough, affecting technologies such as laser-based techniques (Duran et al., 2002). The quality of sewer inspections can therefore be optimised by choosing the most suitable method for the particular situation.

Six hardware inspection methods are briefly covered here: CCTV and optical method methods, acoustic sensor techniques, sonar techniques, infrared thermography, laser profiling and ground penetrating radar. These methods include popular, widely used techniques as well as emerging technologies. Table 2-1 summarises the main characteristics of these various techniques, covering the advantages and limitations inherent in each system.

Traditional CCTV inspection (Duran et al., 2002) remains the most well-established and widely used industry standard inspection technique worldwide. According to a survey published by Thomson & Grada (2004) 100% of participants from large wastewater utilities utilised CCTV as their primary inspection method. CCTV, and other optical based methods such as sewer scanner evaluation technology (SSET) (Civil Engineering Research Foundation, 2001), produce video records or digital images to identify blockages and other failures. The current standard approach employs a remote camera mounted on a motorised crawler controlled by trained operators who examine the regions inside the pipe, and locate and identify the type and severity of any observed defects (Halfawy & Hengmeechai, 2014).

Acoustic monitors (Khan & Patil, 2018) and laser profiling systems (Stanić et al., 2017) are also available commercially. Acoustic technologies, such as pulse reflectometry and sonar, use measuring devices to detect vibrations and/or sound waves to identify impediments or discontinuities caused by blockages and other faults in pipelines. Laser profiling (also known as lightlining) uses a laser to project light onto the sewer pipe wall, indicating changes to the pipe's shape, which can be caused by deformation, corrosion, or siltation.

Table 2-1 Main characteristic of hardware-based blockage detection and location techniques.

.

Technology	Defects Detected	Time Required	Cost	Commercially Available	Advantages	Disadvantages
CCTV	Blockages, Sediment, Roots, Pipe sags, Off-set joints, Cracks, Leaks, Service connections	High	Medium	Yes	Standard Technology Effective Permanent video record Reliable	Slow Labour intensive Prone to error Only inspects pipe above waterline
Acoustic Sensors	Blockages Leaks	High	Medium	Yes	Useful as pre-screening before more detailed inspection	Only detects general distress
Sonar / Ultrasonic	Blockages, Sediment, Roots, Pipe sags, Corrosion, Cracks	High	High	Yes	Suitable for pipes of any material and diameter	Only inspects pipe below waterline
Infra-Red Thermography	Deteriorated pipelines, Leaks, Voids	High	Medium	No	Some forms can be executed from ground surface	Not designed for detection of actual blockages
Laser Profiling	Blockages, Sediment, Roots, Pipe sags, Corrosion,	High	High	Yes	Provides better data quality than CCTV alone Can create 3D models	Only inspects pipe above waterline
Ground Penetrating Radar	Leaks, Voids, Bedding condition	Medium	High	No	Reliable for large non-metallic diameter pipelines	Dependant on operator's experience Impractical in certain soils, Affected by anomalies

There are also a number of emerging techniques such as Infra-red thermography (Huynh et al., 2016) and ground penetrating radar (Ékes et al., 2011). These techniques have been demonstrated in the literature to have potential for sewer pipe blockage detection. However, they have had limited application so far and no commercial systems are available yet.

Finally, the application of robotics in wastewater systems is a relatively new area of research, however robots have been demonstrated to be an extremely promising tool for blockage detection and sewer network management in general (Bischoff & Guhl, 2010). These robots are often equipped with multiple sensors such as sonar, laser profiling, poisonous gas detectors, and can access hazardous environments inaccessible to humans (e.g. IBAK, n.d.; RedZone Robotics, n.d.; Saenz et al., 2010; Streich & Adria, 2004)

For more information on these hardware based technologies a very comprehensive report on techniques for the assessment of wastewater systems is presented by Tuccillo et al. (2009), although it is now a little out of date.

An important factor to consider when considering these hardware techniques is that, given the size of the UK sewer network, only a small sample of the network can be inspected regularly using these systems, especially when taking into account the cost and disruption of accessing underground pipes. Therefore, these methods are generally limited to regular inspections performed to assess the condition of the wastewater pipes. This is beneficial as proactive maintenance can be applied when faults are detected. However, not all blockages are due to deteriorated pipes in bad condition; sewers in good service can also experience blockage failures and so cannot be managed using these techniques (Jin & Mukherjee, 2010). Therefore, wastewater utilities require new technologies which can alert them to the presence of blockages event as occur as they occur. Methods which have the ability to continuously monitor the system have the potential to identify blockages in real time and so prevent imminent failures such as flooding, pollution etc.

2.6.2 Software Based Blockage Detection Techniques

Unlike hardware-based techniques there are few software-based methodologies developed specifically for blockage detection in sewer networks, and none are yet used commonly by wastewater utilities. Two main techniques have been identified: transient analysis and sedimentation modelling. Table 2-2 provides a summary of these methods and their main characteristics. The table also includes a review of statistical models and data mining techniques which are covered in Section 2.6.3.

Software based blockage detection falls under the category of Fault Detection and Diagnosis (FDD), a subfield of control engineering. Faults are generally defined as “an unpermitted deviation or abnormality of at least one characteristic property (feature) of a system from the acceptable, usual or standard range” (Isermann, 1984). FDD aims at identifying faulty elements in a system and diagnosing the cause, so that they can be isolated and repaired and the system brought back to a normal, safe-operating state. This is accomplished by continuously monitoring the system, using FDD to detect and diagnose abnormal conditions and the faults associated with them, then evaluating the significance of the detected faults, and deciding how to respond. Fault diagnosis is an important and active area of research and attracts considerable interest from industrial as well as academic researchers. Early and reliable FDD, while a system is still operating in a controllable region, is highly desirable as it helps to avoid additional damage and reduce any loss of service. This is critical in the correct management of sewer blockages – proactive removal of blockages can prevent overflows and flooding, which are harmful to both the environment and the public.

Following Katipamula & Brambley (2005a, 2005b), automated FDD can be divided into four distinct functional processes, as presented in Figure 2-3. Process 1 monitors the physical system and detects any faults which have occurred. This is generally accomplished by applying one of two methods; either a model is used to calculate the expected values for a system. If the discrepancy or ‘residual’ between the calculated and the measured values is greater than a defined threshold then a fault is said to have occurred. Alternatively, the sensor

Table 2-2 Main characteristics of software-based blockage detection and location techniques.

Method	Inspection/ Monitoring/ Prediction	Cost	Advantages	Disadvantages
Transient Analysis	Inspection	High	• Requires measurements at only 1 location • Non-invasive	• Relies on accurate transients • Long computational times Involves manual & resource intensive processes for data logging and collection
Sedimentation Modelling	Assessment	Low	• Can identify locations most in need of cleaning	• Often has many uncertainties • Often computationally expensive • Cannot identify individual blockage events
Statistical Models & Data Mining Techniques	Assessment	Low	• Not computationally expensive	• Requires large datasets • Cannot identify individual blockage events

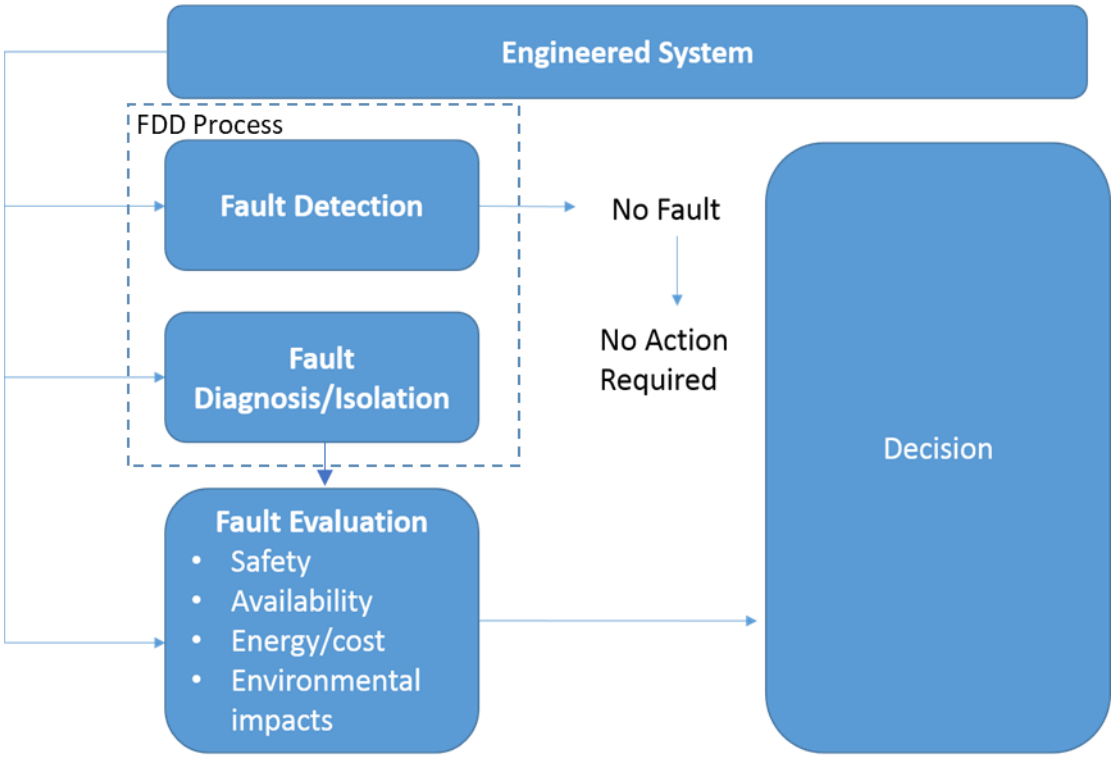


Figure 2-3 Generic application of fault detection and diagnosis to operation and maintenance. Adapted from Katipamula & Brambley (2005b).

readings of the process are monitored, and pattern recognition is used to identify the occurrence of a fault. Process 2 evaluates the fault to determine the cause. Process 3 then evaluates the fault, analysing the severity using factors such as safety, energy use and cost. Finally process 4 makes a decision regarding the correct response to the fault, based on the previous evaluation. Some methods run the detection process continuously and the diagnostic system is activated only when a fault is detected. Other methods run both systems in parallel or run the detection and diagnostic systems in a single step.

FDD has advanced significantly in recent years, primarily due to the major improvements in computer systems. Low level processes which were previously performed by human operators are now automated and performed by computer systems (Venkatasubramanian & Rengaswamy, 2003). However, in general, when abnormal events or emergencies occur, human operators are still required to respond to the problem. This is due primarily to size and complexity of most modern engineering systems and the large variety of malfunctions and diagnostic

procedures which can occur. The major objective in control-engineering is to be able to automate FDD processes.

Few methodologies have been developed with the express purpose of detecting blockage events in sewer systems in real-time so far (excluding methods such as those designed for automated detection of blockages from CCTV images, which can be applied in real-time .(e.g. Kumar et al., 2020). As explained above techniques like CCTV inspections are applied intermittently to portions of the wastewater network and do not operate continually). Hardware based methods are still the most widely used approach by utilities.

Transient Analysis

Transient flow occurs in pipelines due to sudden changes in conditions caused by rapid pressure or flow fluctuations. Pipe blockages, as well as other faults such as leaks and valve operations, generate transient waves which propagate away from the source. Transient-based analysis involves injecting a transient signal into a pipeline and measuring the response at specified locations. The waves are modified by any blockages they encounter and when analysed contain information regarding the properties of the pipe. Transients in water distribution systems for detecting leaks are reasonably well understood. Transients for blockage detection in pipelines are less well studied, however there are a small number of studies in the literature (e.g. Massari et al., 2014; Meniconi et al., 2009; Wang et al., 2005).

Sedimentation modelling

Modelling of sedimentation in sewerage pipes is performed separately from modelling of other blockage types as the mechanisms which result in sedimentation formation is distinct from other types of obstructions (snagging of debris, FOG, pipe collapses etc) and sediment manifests differently in the sewer pipe. Sediment forecasting is a challenging task as sediment deposition exhibits complex physical mechanisms and is influenced by different environmental factors relating to sediment, flow, and channel characteristics. The earliest calculations of sediment transport in steady flow were performed using traditional formulas derived for fluvial flows, for examples see Arthur et al. (1999), May (1982), Nalluri et al. (1994, 1997) and Skipworth et al. (2000). With improving

computational resources and advances in numerical techniques, computational fluid dynamic (CFD) methods have been utilised for sedimentation modelling more recently, e.g. Murali et al. (2020; 2019) and Stovin & Saul (2000). Lately artificial intelligence techniques have also been applied to sedimentation modelling, such as ANNs (Ciğizoğlu, 2002; Kisi, 2005), and fuzzy logic (Azamathulla et al., 2012)

2.6.3 Statistical Models and Data Mining Techniques

Much of the literature on sewer blockage modelling in recent years has focused on understanding and predicting the risk of blockage events, structural defects and other sewer failures and representing their spatial distribution in the wastewater system using statistical and data mining techniques. Although these methods cannot detect the location of individual blockages, they are extremely useful as a risk analysis tool for proactive management of events. Resources including funding, personnel, and access to assets are often constrained and frequent sewer network inspections, e.g. via CCTV, is time consuming and expensive due to the large number of pipes in the wastewater network. Statistical models can be used to predict which sections of the network are most likely to fail, enabling utilities to prioritise these areas for manual inspection, maintenance, and rehabilitation.

Traditionally asset management focused only on detecting blockage hotspots using past blockage incident data. Statistical and data mining techniques, however, can identify the key network parameters which influence the likelihood of a blockage occurring. The information obtained can then be used to formulate proactive blockage maintenance in areas where it is likely that future blockages will form. This is very useful in identifying areas where there is a high risk of blockages in sewers with no recorded history, for example in newly built networks (Arthur & Burkhard, 2010). These techniques are enabled by the ongoing improvement in the quality of sewer asset data and blockage data collected by wastewater utilities, and the ability to link across internal company and external datasets.

These statistical models have indicated that a large range of different factors contribute to the formation of blockage events. There is no overall consensus on

which physical sewer properties can be considered the most influential. For example one of the first sewer blockage studies conducted by Fenner (2000) found that urban catchments containing brick sewers and long pipes with small diameters, shallow depths and moderate to slack gradients experience more frequent blockages. This was supported by Ugarelli et al. (2009) in a statistical analysis of historic data from a sewer network in Oslo. The study also indicated that the age of the sewer and the function (i.e. used for sewage, stormwater or for combined wastewater), has a significant influence on the number of blockages. Another study by Ugarelli et al. (2010) found that slope did not have a great effect on blockage occurrence, although this may have been due to a lack of data. Arthur et al. (2008) found that the risk of blockages was related to predicted locations of flooding, low pipe velocities and smaller sewers and that combined systems are 2.5 times more likely to become blocked compared to separate sewers. Hafskjold et al (2004) observed a similar relationship based on a study in Norway. Marlow et al. (2011) used a web-based survey to collate expert opinion on factors influencing blockage rate from 21 Australian water utilities. Utilities considered drought, sewer attributes, tree coverage, climate and tree planting policy as the most important factors. However, a detailed analysis on blockage data from two water companies showed that the number of blockages was caused primarily by pipe age, pipe diameter and water consumption rates.

One growing field of research in this area in the past two decades has been the use of artificial intelligence methods. For example Bailey et al. (2015) applied decision trees to 8 years of data from the Dŵr Cymru Welsh Water network, finding that smaller, older and shorter sewers in areas of higher property density had a higher rate of blockages. A number of studies have made use of the Evolutionary Polynomial Regression (EPR) model, a hybrid data-mining technique based on evolutionary computing, e.g. Savić et al. (2006). Developed by Giustolisi and Savic (2006) the EPR combines genetic algorithms with numerical regression to develop symbolic models.

Overall, these statistical and data mining methods gave shown that different factors influence blockage formation, although, it has not been possible to identify a clear set of explanatory variables. The relationship between blockages and triggering factors is complex, blockages often appear random and the differences in blockage rates between catchments are difficult to understand. Pipe diameter,

and age have been most consistently identified as having an important affect, particularly if the pipe is over 30 years old. The information provided by these studies is extremely valuable in aiding sewerage asset managers focus proactive maintenance activities, and inspection and replacement programs, by identifying where additional future blockages are likely to occur.

2.7 Fault Detection in Water Distribution Systems

Water Distribution Systems (WDS) experience many problems similar to those which occur in wastewater networks. Although blockages like those found in sewer pipes are not common, many other abnormal events can occur, including bursts, leaks, cracks, corrosion, off-set joints, bedding voids, root intrusion, sediment and pipe sagging. These events can have harmful environmental impacts, cause economic losses, and also have a damaging effect on a water company's operational performance, customer service and reputation.

Leakages in WDS in particular are an important issue and have received significant interest due to the financial cost they place on utilities, as well as the added load on the environment due to wasted energy. The amount of water lost due to leakage in most countries is approximately 20% - 30%. As with sewer pipe blockages, leakage, pipe bursts and other abnormal events are often brought to the attention of water companies only when a member of the public calls in to report a visible event. The water distribution network is a large and complex system, and it is often difficult to detect faults promptly when they first occur. Many countries have developed performance indicators or policies to measure and regulate leakage and water authorities can face strict penalties for failing to address leakage in their network. This has provided the necessary incentives for investment in better leak detection technology and enhanced leak reduction strategies.

As with wastewater networks, the quantity and complexity of network sensors in the WDS is growing at an increasing rate. Smart water networks offer the potential to identify leaks early, thus reducing the amount of water that is wasted and saving utilities money. These solutions involve the use of flow and pressure sensors to gather real-time data, which is then analysed using algorithms to

detect patterns that could reveal a leak in the network and provide information on the location of the leak.

Just as with wastewater blockage detection, leak detection in wastewater distribution can be classified into hardware-based methods and software-based methods (Yuan et al., 2015). A review of the literature reveals that there have been a large number of software-based models and techniques developed for detecting abnormal events in WDS, often using methods which have not yet been applied to wastewater systems. In many cases these have the potential to be utilised in wastewater blockage detection in the future.

Izquierdo (2004) provides a comprehensive review of mathematical models and methods used in the water industry. Early leakage detection models formulated the situation as an optimization problem for computing the magnitudes and locations of leakages, based on flow and pressure measurements. Recently, computational intelligence techniques have gained prominence as tools for leakage and abnormal event detection. A great deal of research has been conducted into data-driven approaches to detect leakage and many different techniques have been applied. A recent review of data-driven approaches for leak detection has been presented by Hu et al. (2021).

In particular many ANN models have been demonstrated to successfully detect leakage. For example, Mounce et al. (2010) developed a Mixture Density Network (MDN) ANN to detect leaks and bursts at the district meter area (DMA). Romano et al. (2012) developed a fully automated Event Recognition System combining several self-learning AI techniques and statistical data analysis tools including wavelets, ANNs, statistical process control techniques, and a Bayesian inference system to infer the probability that a burst has occurred. The ERS was further enhanced by Romano et al. (2013), who included an EA optimisation strategy.

2.8 Industry Practice

This final section of the literature review address industry practice. As stated, wastewater companies have historically relied primarily on hardware-based approaches, and on their customers to report blockage events in the sewage systems. This tends to result in service interruptions and increased customer

complaints, which in turn affects regulatory performance. This is still often the main approach. However, utilities are beginning to develop other, more proactive, techniques. This section presents an overview of blockage detection systems available commercially and an overview of projects implemented by various wastewater utilities.

There are a small number of blockage detection systems available commercially, however there is little information on the details of the systems' methodologies. Environmental Monitoring Solutions provide services such as catchment-based monitoring, flow and load surveys for sewer flows and wastewater treatment works, and hydrometry using Artificial intelligence-based modelling. One of their products is *SMART Sewer* (EMS, 2020), an on-line real-time wastewater network monitoring system designed to identify blockages before they reach a critical level and to differentiate between other causes of level increase. Monitored manholes subdivided into 'triplets' and AI techniques are used to detect the formation of blockages, estimating three probability grades: clean, partially blocked and blocked.

InfoSewer (Innovyze, n.d.-a) is a 1D sewer modelling software developed by Innovyze for use in the planning, design, analysis, and expansion of sanitary, storm and combined sewer collection systems. The software can simulate the transport and settling of sediments. Innovyze also produce *InfoSWMM* (Innovyze, n.d.-b) which can be used for real-time control of hydraulic structures.

Another example is *Meniscus* (Meniscus, n.d.) who provide CSO monitoring and analysis. Their system uses machine learning to compare the historic performance of assets with historic rainfall. This is used to then predict the operation of the CSOs under current and forecast rainfall, providing a range of alerts such as dry weather alerts and post storm alerts (if the CSO is spilling for longer than expected once a rain event has finished). The system also aims to detect partial blockage events, by analysing the rate of change in level in dry weather flow conditions and comparing the current CSO level with data in the upstream CSO.

Nuron are currently developing a fibre technology which they claim act like the human nervous system for sewer networks (Nuron, n.d.-a). The fibre sensors extend along the bottom of sewer pipes and measure flow, depth, temperature,

and structural integrity every five metres along the sewer pipe. This enables potential incidents to be accurately diagnosed, rapidly localised and thus averted. The company are reported to be carrying out a pilot project in conjunction with Northumbrian Water (Nuron, n.d.-b).

Wastewater utilities have also implemented their own projects designed to address CSO pollution incidences. Dŵr Cymru Welsh Water addressed capacity problems in their sewer network. Localised sewer flooding and excessive CSOs were polluting a protected shellfish water and a special area of conservation, resulting in the threat of European Commission Infraction Proceedings. Using a hydraulic model built for the entire sewerage network a range of approaches to decrease CSO spills was designed, including implementing smarter flow control so that existing storage in the system was better utilised (Environment Agency, 2013).

Yorkshire water commissioned the monitoring of high-risk CSOs between January and December 2004 in order to investigate the reasons for investing in long-term data monitoring (Grandison, 2005). The study found important benefits for long-term asset performance management. As a result, a two part £3.2m, pollution prevention programme was implemented in 2006 targeting known ‘hotspots’ across the network, which involved the installation of Hawkeye sensors in CSOs to supply real-time data.

These projects demonstrate that wastewater utilities are beginning to turn towards a more data-driven, proactive approach to blockage management, rather than relying on traditional reactive techniques. However, these proactive methods are not yet widespread, and there are currently only a small number of data-driven blockage detection systems commercially available.

2.9 Summary and Key Gaps in Knowledge

In this chapter a review of the literature relating to the detection and management of blockages events in wastewater networks was presented, in addition to a review of the management and modelling of combined sewer overflows. First, an overview of sewer blockages and CSOs was presented in Sections 2.2 and 2.3. Next a discussion on the increasing availability of sewer network data and rainfall

data was presented in Section 2.4. Section 2.5 presented a review of techniques developed to model CSOs, covering physical and data driven modelling. Section 2.6 then reviewed the available blockage detection techniques, describing hardware and software methods as well as statistical and data mining techniques. An evaluation of fault detection techniques utilised in the water distribution system, which have the potential to be applied to sewer network blockage detection, was then given in Section 2.7. Finally 2.8 described blockage detection systems available commercially, and also covered a number projects implemented by various wastewater utilities in the UK relating to blockage detection and CSO management.

From this review it is evident that a variety of different methods and techniques have been developed for blockage detection, and fault detection in the water sector in general, ranging from hardware-based techniques to hydraulic models to artificial intelligence techniques. No single method for blockage detection is sufficient by itself, and hence techniques are generally used in conjunction with each other. Traditional hardware techniques are still the most popular practice, and there have been many technological advances in recent years, making them more effective and affordable. However, these hardware techniques are generally still costly, invasive, and time-consuming. Many studies have highlighted the benefits of proactive sewer management and techniques such as data-mining and statistical based blockage prediction are being increasingly applied to better direct proactive sewer network maintenance. Nonetheless, despite these developments blockages are still often not identified or removed before they cause serious incidences in the wastewater network. Consequently hundreds of thousands of flooding, pollution and loss of service events occur every year in the UK. As such wastewater utilities still frequently rely on customer complaints to report such events – thus damaging their reputation and increasing the number of customer complaints they receive.

Given the above it can be concluded that there is a significant lack of methods available for real-time blockage detection, i.e. systems that can alert utilities to the presence of blockage events as soon as they manifest, facilitating prompt removal. With the increasing number of sensors installed in the wastewater network in recent years (in the UK this has been facilitated significantly by the implementation of the EDM project), large volumes of sewer level measurements

are now generated in real time. Thus, the development of real-time sewer network models is feasible.

Indeed many studies have concluded that ‘smart draining systems’ will play an important role in future urban drainage management (Edmondson et al., 2018; Lund et al., 2018; Shahra et al., 2019). The increasing implementation of real time control systems in sewer networks demonstrates that this sewer data can be used successfully to better manage the wastewater network in near real-time and, in addition, that wastewater utilities are willing to implement these kinds of systems. Similarly, the implementation of real-time leak detection methodologies for water distribution networks demonstrates that real-time fault detection is a viable and efficient strategy. These leak detection systems have been shown to be very successful in enabling water companies to react promptly to leaks and bursts.

Additionally, this review indicates that data-driven modelling and computational intelligence have demonstrated their applicability to wastewater network modelling. Data driven models are successful at learning the complex nonlinear dynamics of sewer systems. Additionally, the large amount off telemetry data supports a data-driven approach. There are a number of physical-based models designed to simulate sewer flow in wastewater systems, such as the virtual tank model. However, these models are generally computationally expensive and require large amounts of data for calibration.

With above gaps in knowledge in mind, the development and evaluation of a novel data-driven sewer blockage detection methodology is presented in the following chapters. The system is designed to enable a more proactive approach to operation of the sewerage network, facilitating the reduction of flooding and pollution incidents.

Chapter 3: CSO Level Prediction Methodology

3.1 Introduction

In recent years there has been increasing regulatory pressure on wastewater utilities to reduce the number of unconsented spills that occur in their networks. With predicted changes to rainfall patterns and intensity due to climate change and increasing urbanisation, the resilience of wastewater networks against flooding is a growing concern.

Successful management of sewer systems requires a near real-time understanding of network behaviour. If utilities have the capacity to predict discharge events before they occur preventative measures (e.g. maximizing storage, adjusting set points for movable gates and pumping stations) can be implemented to mitigate their effects (Morales et al., 2017) and warnings can be issued to the public for unavoidable spills.

Advances in smart infrastructure and the emergence of the Internet of Things (IoT) presents the opportunity to exploit sewer network data for operational management using data-driven techniques. CSO level sensors are being increasingly installed in sewer systems throughout the UK and worldwide, providing useful, real time information.

This chapter presents a methodology developed to forecast levels in a CSO chamber up to six hours ahead, utilising CSO level data and rainfall data. The methodology employs an artificial neural network (ANN) model and is designed to operate in real time, providing an early warning for upcoming spills events. The forecast results are also used in the event detection system, described in Chapter 4.

Three different ANN models have been developed: (i) a simple artificial neural network model, (ii) an evolutionary artificial neural network (EANN) model and (iii) a novel bi-model committee evolutionary artificial neural network (CEANN).

The results from the different models are compared and contrasted in the Chapter 5.

3.2 Data Pre-Processing

Data pre-processing is an important process in all data-driven modelling. Both rainfall and CSO level data may contain erroneous or missing data-points, which can affect and have a significant impact on the quality of the forecast models. Pre-processing of the data is performed to eliminate errors and ensure a continuous, uniform, and filtered data stream. It is necessary here to distinguish erroneous data from any genuine outliers, for example those caused by extreme rainfall events or due to blockages and other failures in the sewer network.

The first step in the data pre-processing methodology consists of checking and correcting any erroneous timesteps in the data caused by probe/equipment malfunctions. For both rainfall and level data streams duplicate datapoints are removed and any negative values are removed. For level data only, values above the height of the CSO level monitor are removed, providing that the type of monitor used cannot measure above its installation height. This process is performed on both historic datasets and real-time data.

3.2.1 Benching Removal Methodology

The second step in the pre-processing is benching identification and removal. ‘Benching’ refers to false data readings caused by the level monitor in the sewer mistakenly measuring the height of objects in the CSO chamber instead of the actual water level. These objects include, for example, the bracket holding the level monitor or the wall of the CSO chamber. An example of typical CSO level data containing benching is shown in Figure 3-1.

Not only does the occurrence of benching significantly affect data quality, benching which occurs above the spill level of the CSO chamber (as it does in Figure 3-1) can cause the CSO to mistakenly appear to be frequently spilling out of consent. As wastewater utilities are required to report overflow frequencies and duration to the environment agency (Environment Agency, 2018a) it is extremely

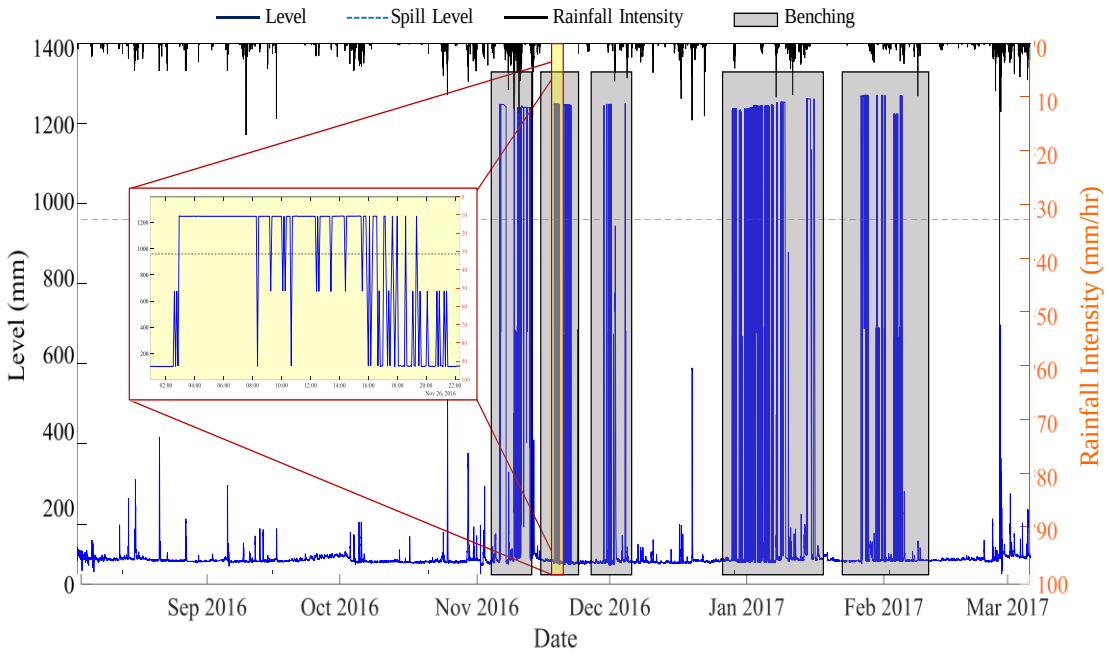


Figure 3-1 CSO chamber level data containing benching.

beneficial for utilities to be able to effectively identify and remove these false readings.

The occurrence of benching is relatively common in CSO level data, an analysis of 100 monitors installed in CSO chambers in the United Utilities network indicated that 36 suffered from some degree of benching.

Benching is characterised by sharp peaks in the CSO level data, unrelated to rainfall or the usual diurnal CSO pattern. These peaks can be distinguished from genuine increases in level caused by rainfall or events such as blockages as they generally occur progressively over multiple timesteps. As benching is caused by the presence of objects in the CSO chamber these incorrect data points largely occur at the same level, within a range of a few centimetres. In some cases benching is caused by multiple objects in the chamber, resulting in benching occurring on several levels, an example of this is illustrated in Figure 3-2.

There are numerous methods available for data cleansing, however, none are designed for the unique profile exhibited by benching. Thus, a methodology has been developed specifically to identify and remove benching in both historic and real time level data, presented below.

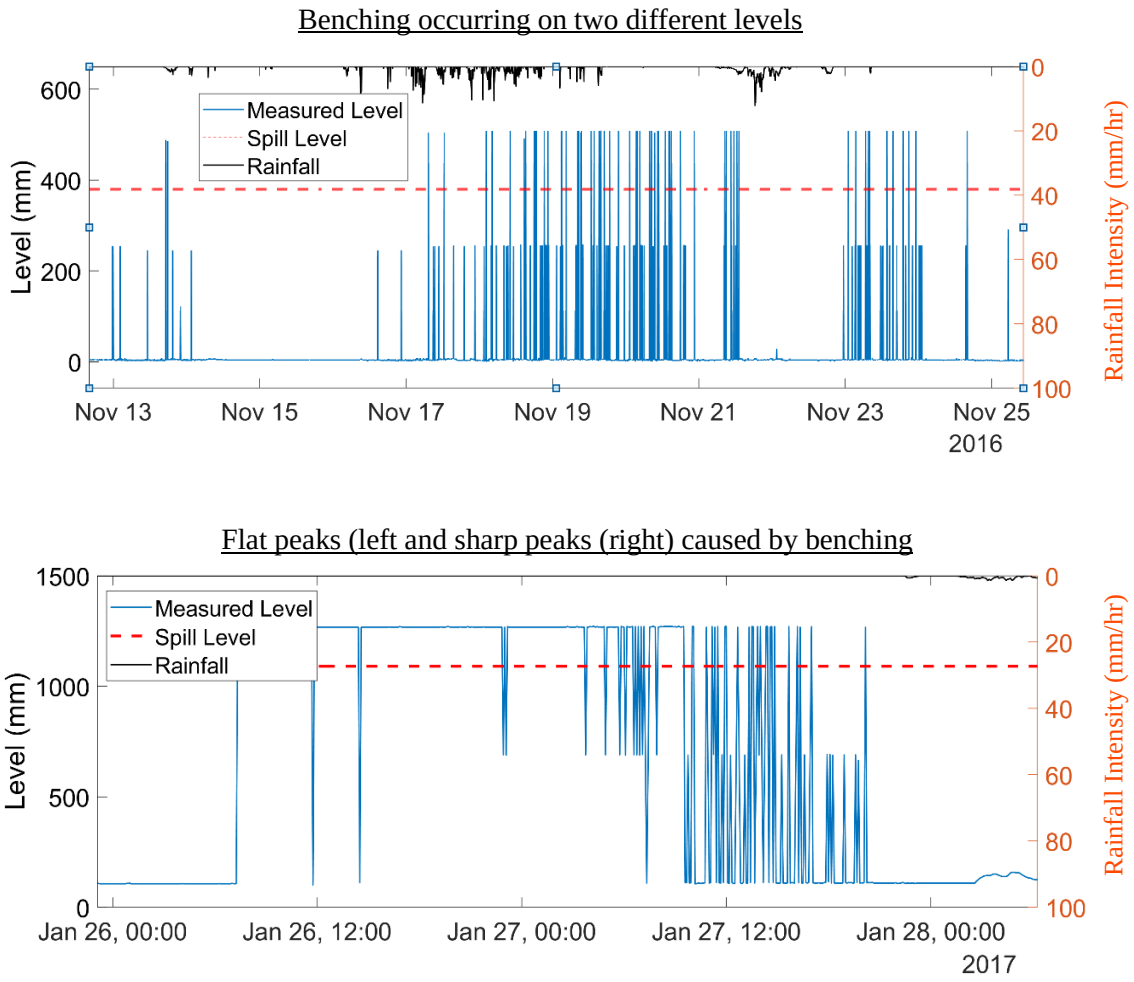


Figure 3-2 Examples of different types of benching in CSO level data.

3.2.2 Historic Benching Removal

The process for removing benching from historic level datasets is as follows:

Step 1: Extraction of dry weather level data.

Dry weather data refers to time periods without rainfall. Analysing only dry weather periods excludes any increases in CSO level caused by rainfall events, thus making it easier to identify level fluctuations due to benching. Note that the majority of the data occurs during dry weather (~98%), therefore excluding wet weather periods does not remove a significant portion of any potential benching present in the data. The method utilised to extricate the dry and wet weather time periods is explained in further detail in Section 3.6.

Step 2: Identification of local maxima

Next local maxima in the dry weather data are identified. Local maxima refer to data points significantly larger than their neighbouring values i.e. sharp peaks that immediately return to the previous ‘normal’ level. These peaks are defined here as data points greater than their neighbours by a minimum of 30% of the distance between the average dry weather CSO level and the spill height of the CSO chamber. As benching is generally characterised as occurring over a single data point only, these maxima are restricted to a maximum peak width of two timesteps.

Step 3: Identification of flat peaks

In addition to the typical sharp peaks (displayed in Figure 3-1), benching may also manifest as consecutive points of the same value (within a few millimetre tolerance), i.e. flat peaks, as can be seen in Figure 3-2. As this signature does not occur naturally in genuine level data these points are unambiguously erroneous. Flat peaks can also be a result of logger malfunction. However, flat peaks due to benching can be identified as they occur at the same height as sharp spikes in the data also caused by benching.

Step 4: Peak Analysis

The identified peaks are then analysed to determine if they are caused by benching, or if they are genuine spikes in the level data. As stated, benching at a given site results in erroneous data only at a small number of heights (commonly just 1). Therefore, by determining if the identified peaks (both sharp and flat) occur significantly at particular levels, the presence of benching at a CSO can be confirmed and the instances of benching in the level data identified and subsequently removed.

This is performed by generating a histogram of the peak levels, grouping the data according to height. If the number of data points in a single bin is over a set threshold, then benching at that height is inferred. Based on analysis of CSO data the threshold has been set at 30 data points per year of data when using a histogram bin size of 5 mm.

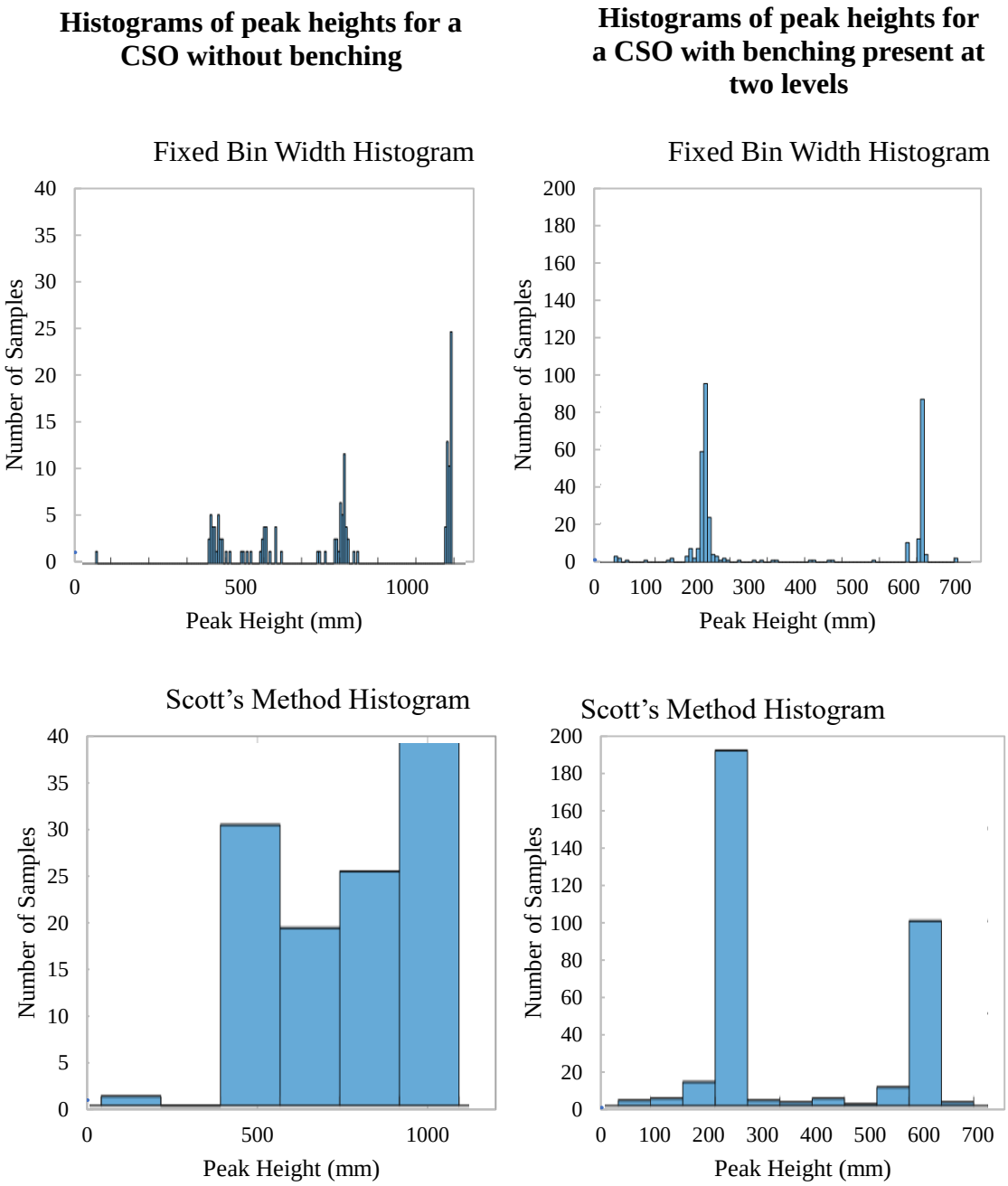


Figure 3-3 Histograms of peak heights using a fixed bin width of 5 mm, vs using Scott's method for determining bin width.

Automatic methods for determining the bin width of the histogram, such as Scott's method (Scott, 1979), were considered. These methods are based on the number of data samples and the variance of the data etc., rather than using a fixed bin width. The benefit of these approaches is that they are adaptive to the data and work well when benching is present as there are a large number of peak samples. In these cases, the peak samples caused by benching tend to be placed in one

bin, which can be clearly identified. However, in some cases these bin widths are relatively large, making it difficult to determine the exact benching height. Additionally, these methods do not work well when applied to sites with no benching present. As these sites have no levels with a large number of samples this results in overly large bin widths being estimated. In some cases the bins are large enough that the number of samples exceeds the benching threshold and benching is falsely detected.

A comparison of the histograms generated using Scott's method and using a fixed bin width of 5 mm is presented in Figure 3-3, when applied to a CSO with benching and a CSO without benching. Note that for the CSO with benching the benching occurs at 2 different levels – the peaks of which can clearly be seen in the histograms. Whereas the histograms of the CSO without benching contain no bins with a substantial number of samples.

Step 5: Removal of Benching Data

If benching is determined to have occurred during Step 4, the identified dry weather peaks at the benching height/s are removed. Additionally, flat and sharp peaks are identified in the wet weather data (using the methodology described in Steps 2 and 3). Any of these peaks occurring at the identified benching height/s are also removed.

3.2.3 Real Time Benching Removal

During real-time running of the methodology the incoming rainfall and CSO level data must be processed in a sequential form as it arrives. Therefore, the benching removal process must also be performed in a real-time fashion. The real-time benching removal methodology is similar to the historic benching removal process. CSO level data is analysed at each timestep. Local maxima in the level data are identified and, if they occur at benching height, are removed. Flat peaks are also identified and, regardless of the height, are removed.

This benching removal process is only conducted on incoming real-time data if the initial analysis of the historic dataset during the model training indicated that benching was present. This is to ensure there is enough data to confidently identify the benching height. As the system is designed to be retrained every few

months to account for changes in the network, any benching which begins in a CSO after the system has begun running in real time will be identified during this period.

3.2.4 Missing Data

Once the erroneous data has been removed, the final processing step involves infilling any missing level and rainfall data. Regarding rainfall data, one or two missing data points per day are not uncommon, however, there are also rarer instances where data is not reported for many hours or even days. Regarding the level data, missing data is less frequent. However, level monitor malfunctions can cause the monitor to stop reporting. In extreme cases normal functioning is only resumed once the utility is alerted to the problem and personnel sent to the site to repair the monitor, potentially resulting in many weeks or even months of missing data. Additionally, level data may also be intentionally removed, as described above due to benching or other malfunctions.

It is important to ensure missing data is only imputed, therefore, if the number of missing values are under a certain level. For level data and rainfall data during rainfall events, missing data less than 6 consecutive datapoints in length are infilled. As mentioned above, large periods of the rainfall dataset are dry weather, i.e. zero, and it is therefore possible to confidently infill larger periods of missing rainfall data during these periods. For periods where the sum of the data points for the 12 hours immediately preceding and following the missing values are zero, missing rainfall data points less than 12 consecutive timesteps in length are also infilled.

There are several different techniques for imputing missing values; selection of the most appropriate method can be a difficult choice. Here, missing data has been estimated using linear interpolation. This technique assumes a linear relationship between data points and utilises non-missing values from adjacent data points to compute a value for a missing datapoint (Lepot et al., 2017). Linear interpolation was selected as it is fast, easy to use, and has been demonstrated to produce good results when imputing time-series data (Norazian et al., 2008). According to Gnauck (2004) linear interpolation is efficient and generally better than non-linear interpolations for predicting missing values in environmental

phenomena with constant rates. Missing datapoints which cannot be generated are infilled with 'Not a Number' to ensure a uniform dataset is constructed.

More confidence in this infilling process could be gained by cross-referencing the rainfall data with the level data, and this approach could be implemented in the future. Additionally, missing radar rainfall data could be infilled using data from neighbouring radar rainfall grid squares. An analysis comparing the relationship between CSO level data and radar rainfall data indicated that level data and radar data from multiple neighbouring grid squares have a strong correlation (see Figure 5-7). Therefore using data from nearby grid squares is likely to provide acceptable results. However, an analysis of pre-processed level and rainfall data found very few instances of this method inaccurately infilling missing datapoints. Thus the current approach was deemed adequate for the case studies represented in this thesis.

3.3 Model Performance Metrics

According to (Legates & McCabe, 1999) to adequately assess hydrologic model performance at least one goodness of fit measure (such as the Nash-Sutcliffe efficiency coefficient (NSE)) and one absolute error measure (such as the mean squared error (MSE)) should be calculated. This study applies 4 model performance metrics to evaluate the performance of the different ANN models: The Nash-Sutcliffe efficiency coefficient, the mean squared error, the mean absolute percentage error (MAPE) and the structural similarity (SSIM) index.

Nash-Sutcliffe Efficiency

The NSE is a normalized statistic that determines the relative magnitude of the residual variance, or 'noise', compared to the measured data variance (Nash and Sutcliffe, 1970). It is defined as

$$NSE = 1 - \frac{\sum_{i=1}^{n}(O_i - P_i)^2}{\sum_{i=1}^{n}(O_i - \overline{P_i})^2}$$

where O and P are the observed and predicted values respectively. The NSE ranges from 1 (a perfect fit) to -∞, with values between 0 and 1 regarded as acceptable, and greater than 0.5 as good. An NSE value below 0 indicates that

the mean value of the observed time series would produce more accurate results than the model.

Along with the MSE, the Nash-Sutcliffe efficiency is perhaps the most widely used statistic for assessing the goodness of fit of hydrologic models. However, it is important to note that the differences between the observed and predicted values are calculated as squared values – thus larger values in the time series are overestimated compared to smaller values. As a result, the NSE may overestimate model performance during peak flow and underestimate performance during low flow.

Mean Absolute Percentage Error

The MAPE is defined as

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{O_i - P_i}{O_i}\right|$$

where n is the total number of timesteps. The MAPE is useful as an absolute measure of forecasting accuracy; as the error is not squared it is less sensitive to large errors that occur at high magnitude. However, it is subject to 'fouling' by small values.

Mean Squared Error

The MSE is a commonly used error measurement in many diverse disciplines. It is defined as

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(O_i - P_i)^2$$

The MSE calculates error in the units of the model, which can be beneficial for analysis. However, the metric penalises models that exhibit large deviations, therefore a few large errors can cause a large MSE error, even where most of the forecasting error are within an acceptable range (Hwang et al., 2012).

Mean Structural Similarity Index

The structural similarity (SSIM) index, also known as the Wang-Bovik index, is a metric originally developed to compare the quality of digital images and videos in

order to quantify image degradation caused by data processing (Wang et al., 2004). The index measures how similar the processed image is from a reference (i.e. perfect) image with respect to the structure within a convolution window. SSIM is a perception-based model (in contrast to an absolute measure like the MSE).

For images X and Y (computed as matrices of pixels) the SSIM between two windows x and y of common size N×N is calculated as:

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)},$$

$$c_1 = (k_1 L)^2 \qquad c_2 = (k_2 L)^2$$

Where *L* is the dynamic range of the image. *k1* and *k2* are arbitrarily taken as 0.01 and 0.03 respectively and are used to ensure that near-zero denominators do not cause computational instability.

The mean SSIM (MSSIM) for the global image similarity is then given as

$$MSSIM = \frac{1}{n}\sum_{j=1}^{n} SSIM(x_j\, y_j)$$

where MSSIM = 1 if the images are identical.

The SSIM was developed for image analysis, however, it can be applied to time-series data by considering the signal as a N×1-pixel image and using a 1-dimensional convolution window over the time series. SSIM has been used successfully for similarity analysis in disciplines such as biological neurogram signals (Hines & Harte, 2012), speech signals (Hines et al., 2012) and aeroacoustics. (Breakey & Meskell, 2013).

The metric has not, to our knowledge, been applied to hydraulic models. However, Mo et al., (2013) stated that the SSIM could have a novel application potential in hydrology and meteorology. Evaluating the applicability of four non-traditional similarity metrics for hydrometeorology data, they concluded that the SSIM has an advantage over the other metrics as it takes into account the pattern correlation between two compared objects, and also uses the differences in the

means and variances to penalize the degree of similarity. Unlike traditional metrics, SSIM index is designed to capture the perceived structural variation rather than the perceived error (Wang et al., 2004).

In this study the MSSIM is calculated for the EANN model analysis by taking *x* and *y* as the observed and forecast values and *L* as the range of the CSO water level.

3.4 Simple Artificial Neural Network Model

The basic ANN model used in this methodology consists of a feed forward ANN with a single hidden layer trained using the backpropagation method (illustrated in Figure 3-4). The back propagation neural network (BPNN), proposed by Rumelhart and McClelland (McClelland et al., 1986) is the most commonly applied ANN structure. Using this method, information is propagated in one direction only, from the input layer to the output layer. The use of a multi-layer perception, i.e. a model containing hidden layers, means that, unlike a single layer perception, the model is able learn non-linear functions. A single hidden

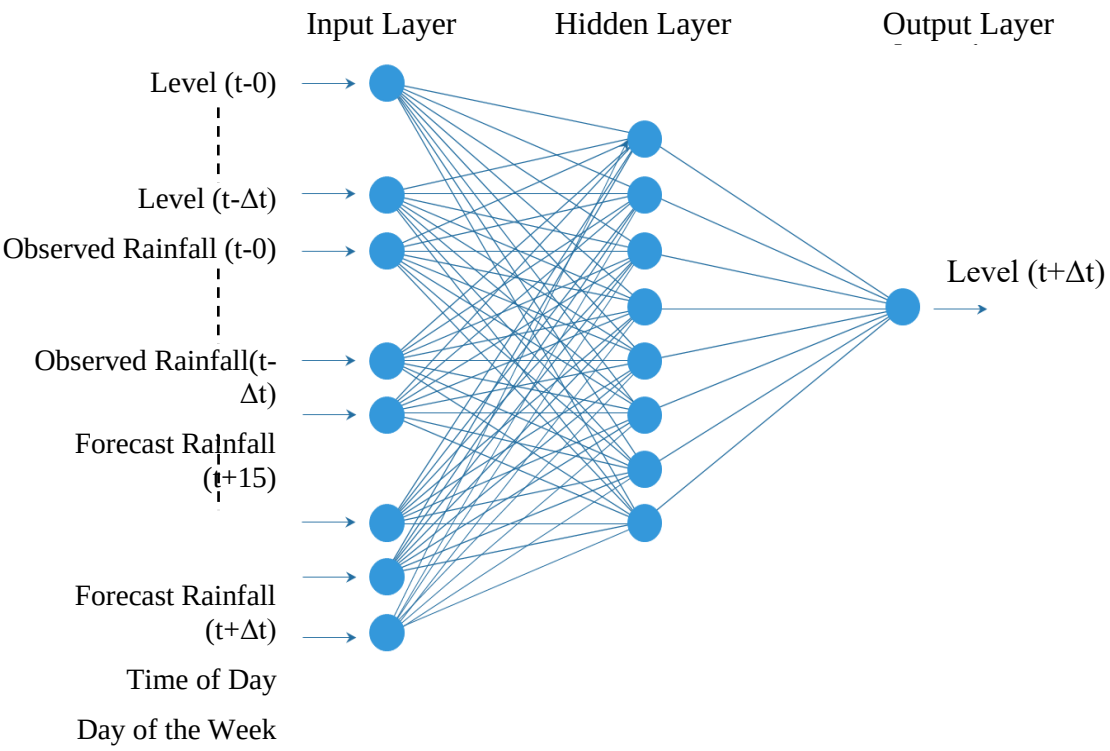


Figure 3-4 Schematic diagram of a generic artificial neural network structure.

layer was selected as they have been demonstrated to be sufficient for the large majority of problems.

According to the universal approximation theorem a feedforward neural network with a single hidden layer containing a finite number of neurons, can approximate any continuous function on compact (closed and bounded) subsets of n-dimensional Euclidean space, under mild assumptions on the activation function (Hornik et al., 1989).

A hyperbolic tangent transfer function is used for neurons in the hidden layer and a linear transfer function for neurons in the output layer. This setup was identified by Romano & Kapelan (2014) as producing accurate results with fast training times. Inputs to the networks consist of antecedent CSO level data, antecedent rainfall data, and forecast rainfall data.

A review of the literature found that forecast rainfall data has not previously been used in ANN CSO level forecasting. Analysis demonstrated that the use of the rainfall forecasts in addition to antecedent rainfall data significantly improves the accuracy and forecast range of level predictions.

However, it is important to note that the accuracy of forecast rainfall data degrades at higher forecast horizons. This has been demonstrated in a number of investigations into MET office nowcast data (Clark et al., 2016; Golding et al., 2014; Simonin et al., 2017). Figure 3-5 illustrates this for a point location in

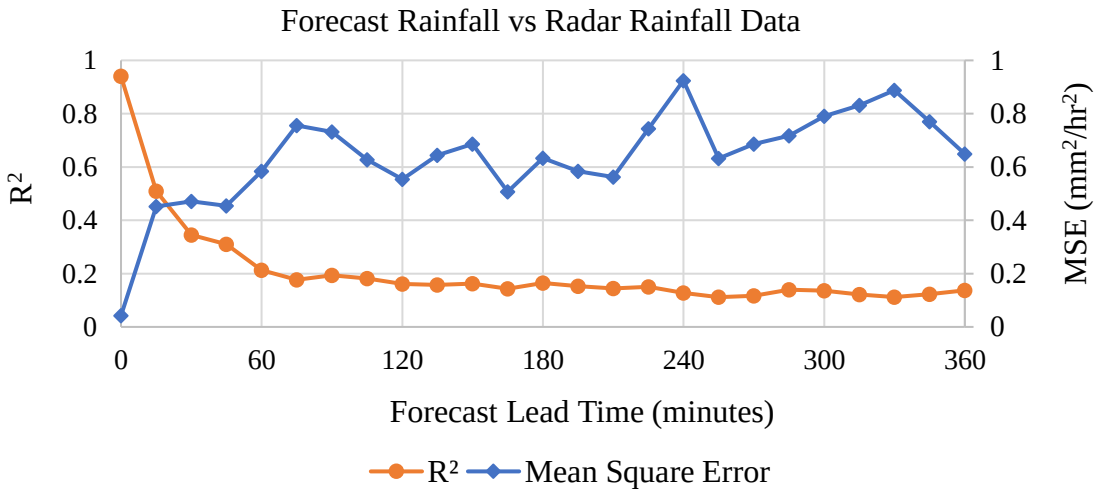


Figure 3-5 R^2 and MSE of forecast rainfall data compared to radar rainfall data as a function of forecast lead time.

Northern England; the graph displays the accuracy of MET office forecast rainfall intensity data when compared radar rainfall data as a function of forecast lead time, measured using the R^2 coefficient and the mean squared error. One year of forecast rainfall data was analysed, from March 2016 to March 2017. As can be seen from this figure, the forecast rainfall data is in good agreement at low lead times, however, accuracy decreases rapidly for forecasts between 0 and 75 minutes, levelling off and decreasing slowly between 75 to 360 minutes. These errors therefore affect the accuracy of the ANN model predictions at larger forecast times.

However, as will be demonstrated in Chapter 5, even at larger forecast intervals, where the R^2 value between forecast and observed radar data is very low, the ANN model with actual forecast data consistently produces superior results compared to the model without rainfall forecasts. Therefore, using forecast rainfall as a model input is beneficial.

When developing ANNs the selection of the model architecture and inputs is of great importance and can have a significant impact on the forecast accuracy. Different CSOs often exhibit very different behaviours and responses to rainfall events. Indeed, even for the same CSO chamber different forecast horizons require different model inputs and structures to produce optimal results. As a result, using pre-defined parameters can result in poor predictions.

This selection process is accomplished by constructing various ANN models with different configurations of parameters and input structures and input structures. The ANN model with the best generalisation capacity is selected by identifying, for each parameter, which value produces the best performance across all 24 forecast horizons (evaluated using the Nash-Sutcliffe efficiency) over the validation period data (i.e. unseen data run as if in real time). This trial and error selection process is performed independently for each new CSO site that the ANN model is applied to, for example if the model is applied to three CSO sites the architecture is selected three times. It is important to note that this process is extremely time consuming and labour intensive and generally evaluates only a small number of all the possible model configurations.

The ANN decision variables and their various ranges of variability presented in Table 3-1. The lag size of the radar rainfall and level data refers to the number of

Table 3-1 Decision variables and associated ranges of variability of the ANN model.

Decision variables	Range of values used in optimisation
Level data lag size (number of time steps)	2 - 72
Radar rainfall data lag size (number of time steps)	2 - 72
Forecast rainfall data lag size (number of time steps)	1 - 24
Time of day	Use/ Do not use
Day of week	Use/ Do not use
Number of hidden neurons	10 - 100
Number of training cycles	50 - 500
Coefficient of weight decay regularisation (α)	0 - 1000

antecedent timesteps used as inputs to the model. Feed-forward ANNs have no internal memory to store past information and as a result cannot process time series data satisfactorily. Thus, a sliding time window approach is employed, whereby past data is input to the network using a window of lagged level and rainfall data. The selection of the data lag size is an important consideration as too small a window and the model may not capture the necessary dynamics of the system, However, too large and the learning time may be prolonged and unnecessary information appear as noise.

Similarly, selection of the number of hidden neurons is an important decision and can significantly affect model performance. A network with too few hidden neurons may have poor accuracy, however an excessive number can decrease the generalisation ability of the model due to overfitting (Sheela & Deepa, 2013).

Overfitting refers to models which correspond too closely or exactly to information contained in the training data and may therefore fail to fit additional data or predict future observations reliably (Hawkins 2004). Rather than learning to generalize from a trend, the overfitting model memorises non-predictive features in the

training data. Overfitting in ANNs is an often-encountered problem and research into techniques to overcome the issue have attracted considerable attention e.g. Piotrowski & Napiorkowski (2013) and Rosin & Fierens (1995).

In addition to selecting an appropriate number of hidden neurons, two of the most commonly applied approaches for avoiding overfitting are also applied here – early stopping and weight decay regulation (WDR). The ranges of variability for these parameters are shown in Table 3-1. Early stopping is the simplest and most widely used method to avoid overfitting, it works by stopping the model training before the training set has been learned perfectly. Selecting the number of training cycles can have a considerable effect on the model performance. The challenge is to train the network long enough that it is capable of learning the mapping from inputs to outputs, but not so long that it overfits the training data. The approach applied here is to treat the number of training cycles as a parameter and select the optimal value. In addition to avoiding overfitting early stopping also often shortens the training time significantly which can be advantageous.

Weight decay regulation is a technique used to prevent overfitting by preventing the network weights from growing too large. Large magnitude weights tend to generate an irregular fit to the data as the model adapts to any noise present. Weight decay regulation penalises large weights by adding a weight decay penalty term at the end of the objective function. The weight decay coefficient is commonly a value between 0 and 1.

Weight decay is one of the most commonly used techniques for regularising parametric machine learning models. The grid search and random search method are the most widely used strategies to select a suitable value for the weight decay coefficient (Bergstra and Bengio 2012). The random search method selects values randomly, whereas the grid search method searches for the best parameter from a manually specified grid of values. However, these methods can be very time consuming (Montavon, Orr, and Müller 2012).

Lastly, the time of the day and the week are included as optional model inputs. This is because water flow in sewers generally exhibit significant hourly and weekday/weekend trends when rainfall is not present. Figure 3-6 displays a typical example of mean daily levels in a CSO chamber for each day of the week

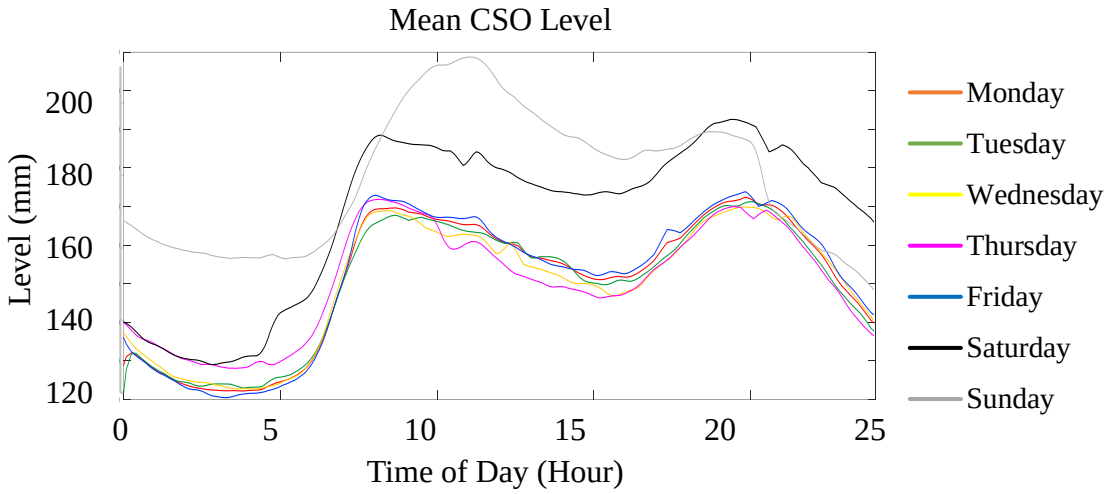


Figure 3-6 Average daily water levels in a CSO chamber located in the United Utilities network for each day of the week.

during dry weather. As can be seen, the weekend levels are significantly higher than the weekday levels, and a strong diurnal pattern is present. To represent this in the model inputs the day of the week and the time of day associated with the forecast horizon, converted into a field representation (i.e. ones and zeros), are potentially used as inputs to the ANN.

3.5 Evolutionary Artificial Neural Network Model

As explained above, when developing ANNs the selection of the various model architecture parameters and inputs described in Table 3-1 is extremely important and can have a significant impact on the model performance.

Trial and error allows the selection of effective variables, however it is a time-consuming, labour-intensive process and requires a human expert. However, when forecasting CSO level on large scale the selection of parameters and input structure manually would not be feasible. When utilised by a real wastewater utility, the methodology could potentially be applied to hundreds of CSO chambers for many different forecasting horizons.

A solution to this problem is to utilise an evolutionary artificial neural network (EANN) model. Information regarding EANNs and their use in hydrology is resented in Section 2.5.2 The main advantages to using an EANN are their adaptability to dynamic environments and the fact that they significantly reduce

the effort required from a human expert to “design” the ANN model for the given situation, whilst also replicating or, as is often the case, outperforming the quality of the results achievable through human expert intervention. Thus, a utility will be able to apply the generic methodology to any CSO chamber equipped with a level sensor and in a location with available rainfall data, without requiring any additional data or expert knowledge. There has been increasing interest in applying EANNs to hydraulic systems in the last decade. However, to our knowledge, EANNs have not yet been applied to CSO level forecasting.

The EANN used here makes use of the basic formulation ANN described in Section 3.4. The EANN employs an Evolutionary Strategy algorithm (Schwefel, 1998) to automatically select the optimal (i.e. that yields the best forecasting performance) ANN input structure and parameter set for the specific training data and forecast horizon considered (in the case studies presented in Chapter 5 24 forecast horizons are considered, from 15 minutes to 6 hours ahead at 15 minute intervals). The EA methodology is based on the design presented by Romano and Kapelan (2014) who developed an adaptive Water Demand Forecasting (WDF) model to predict water demand up to 24 hours ahead.

The EA process is presented in Figure 3-7. The network parameters are initially randomly selected. For each cycle of the evolutionary strategy algorithm the ANN model prediction error (or ‘fitness’) is computed on the testing dataset using the Nash-Sutcliffe efficiency coefficient (McCuen et al., 2006) and used in the objective function of the evolutionary strategy. The fittest networks are allowed to survive while the weak networks are terminated, i.e. the ‘survival of the fittest’, thus increasing the overall fitness of the population. The fit networks are selected to reproduce by generating copies of their genotypes with the addition of changes (or mutations) introduced – these new networks then compete with the parents for a place in the next generation. This process is repeated for a number of generations until a pre-defined termination criterion is satisfied – resulting in the selection of the combination of parameters which produces the lowest error

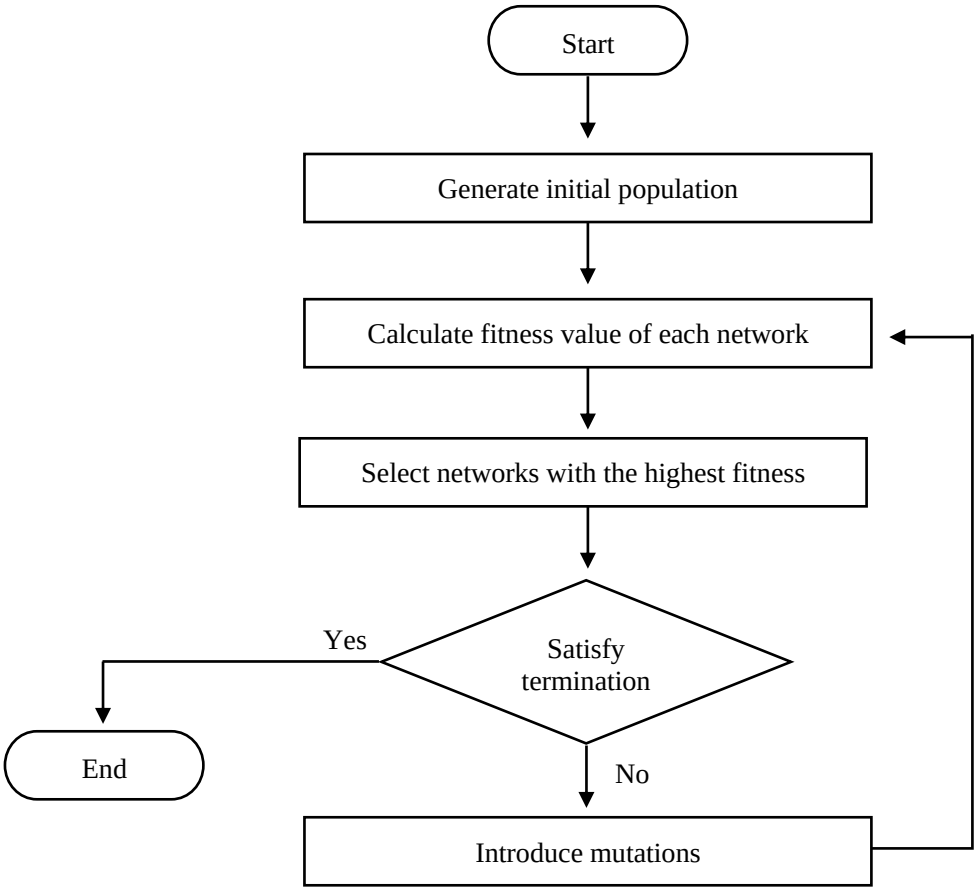


Figure 3-7 Evolutionary algorithm process.

during testing (Benbassat & Sipper, 2013). The evolved parameters are the same as those determined for the simple ANN model, presented in Table 3-1.

The Evolutionary Strategy algorithm parameters used in this work are presented in Table 3-2. The termination criteria is set as the total number of fitness function evaluations ($N_{f.f.e}$) i.e. the number of cycles of the EA run. An isotropic Gaussian mutation operator is employed to introduce the mutations in the child networks. This is a well-known operator which perturbs each component independently using a random number from a Gaussian distribution with zero mean and standard deviation σ. The standard deviation, also known as the step size, determines the strength of the introduced mutation - a large standard deviation increases the exploration of the search space, whilst a small standard deviation promotes exploration of the parent information. Lastly, the selection operator can be set as either '+' or ','. This parameter controls if, as is used in this study, both the parent and child networks are included in the selection pool evaluated to determine the fittest networks (when set as '+'), or if only the newly generated child networks are included in the selection pool (when set as ',').

Table 3-2 Evolutionary strategy algorithm parameters.

Parameter	Value
Number of parents per generation (μ)	10
Number of children per generation (λ)	20
Number of fitness function evaluations ($N_{f.f.e}$)	210
Probability of a parameter being perturbed (P_{mut})	0.6
Standard Deviation (σ)	0.75
Selection Operator	+

The values for the Evolutionary Strategy Algorithm Parameters are based on those identified by (Romano 2012) who developed an EANN to forecast water demand. It was not necessary to conduct a full sensitivity analysis to select optimal values as ES parameters have been demonstrated to be rather insensitive to exact values. According to De Jong, (2007) 'getting "in the ball park" is generally sufficient for good EA performance [and] the EA parameter "sweet spot" is reasonably large and easy to find'. As a result, most EAs today come with a default set of static parameter values that have been found to be quite robust in practice. This was confirmed by Romano (2012) who conducted a sensitivity type analysis and found that changing the EA parameters did not have a significant effect on results and that a large range of EA parameters produced ANN models with very good performance.

3.6 Committee Evolutionary Artificial Neural Network Model

The third and final ANN-based model consists of a Committee Evolutionary Artificial Neural Network (CEANN), also known as a committee machine. The CEANN utilises the EANN model described above, with the same inputs and outputs, and using an EA to optimise the network structure and inputs, and model parameters presented Tables Table 3-1 and Table 3-2 respectively.

A committee machine is a type of neural network which employs the principle of 'divide and conquer'. Rather than using only one ANN model the results from multiple parallel ANNs are combined into a single output, with the aim that the overall result is superior and more robust than any single network acting alone (Tadeusiewicz, 1995). Using this approach committee machines can produce significantly improved results with little extra computational effort.

A committee approach is utilised in this study to overcome data bias caused by imbalanced data. Imbalanced data refers to data sets where the number of samples of one class are significantly fewer than the other classes (Vluymans, 2019). Class imbalance is a serious problem in machine learning. As neural networks and most other machine learning algorithms aim to optimise overall classification accuracy the learning algorithm tends to be biased towards the majority class and the minority class samples are likely to be misclassified. At the same time the minority class is generally more important from a data mining perspective, and despite its rareness may contain important information.

In the case of CSO level modelling a significant majority of rainfall data is dry (i.e. 0 mm/hr), and of the wet weather data the majority is very light (under 1 mm/hr), heavy rainfall events are very infrequent. Figure 3-8 shows the distribution of rainfall intensity from a CSO site in Northern England – the graph shows 2 years of observed 5-minute MET office rainfall data. At this site 89% of data points are

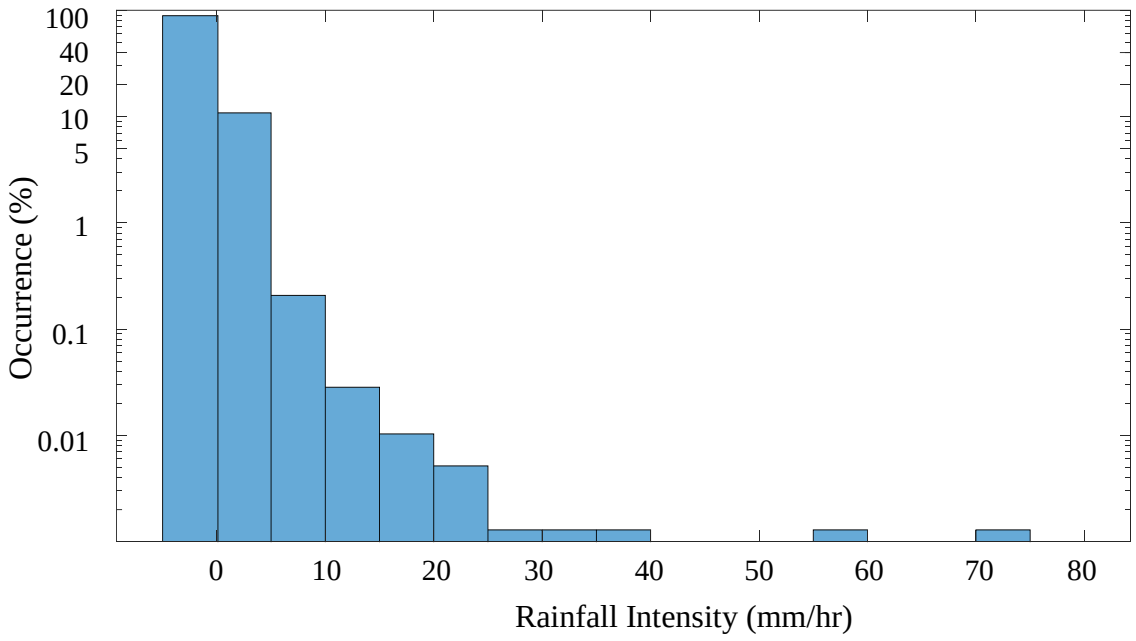


Figure 3-8 Distribution of rainfall intensity data from a CSO site (log scale).

0 mm/hr and 99.6% are under 4 mm/hr (the MET office definition of heavy rainfall (Met Office, 2007)).

Analysis of ANN model forecasts demonstrate that this data imbalance results in good dry weather performance but very poor wet weather performance, especially at high forecast horizons. This can be seen clearly in Figure 3-9, which displays the performance of the EANN model described above, optimised for three different CSO sites. The figure compares the performance of the models on unseen data when forecasting wet weather timesteps only, dry weather timesteps only and for all timesteps. The model performance is measured using the Nash-Sutcliffe efficiency. All three examples show that the forecasts are significantly worse when forecasting wet weather. Note however, that the overall performance of the models over all timesteps is still good, as the wet weather timesteps comprise such as small proportion of the data. Wastewater utilities

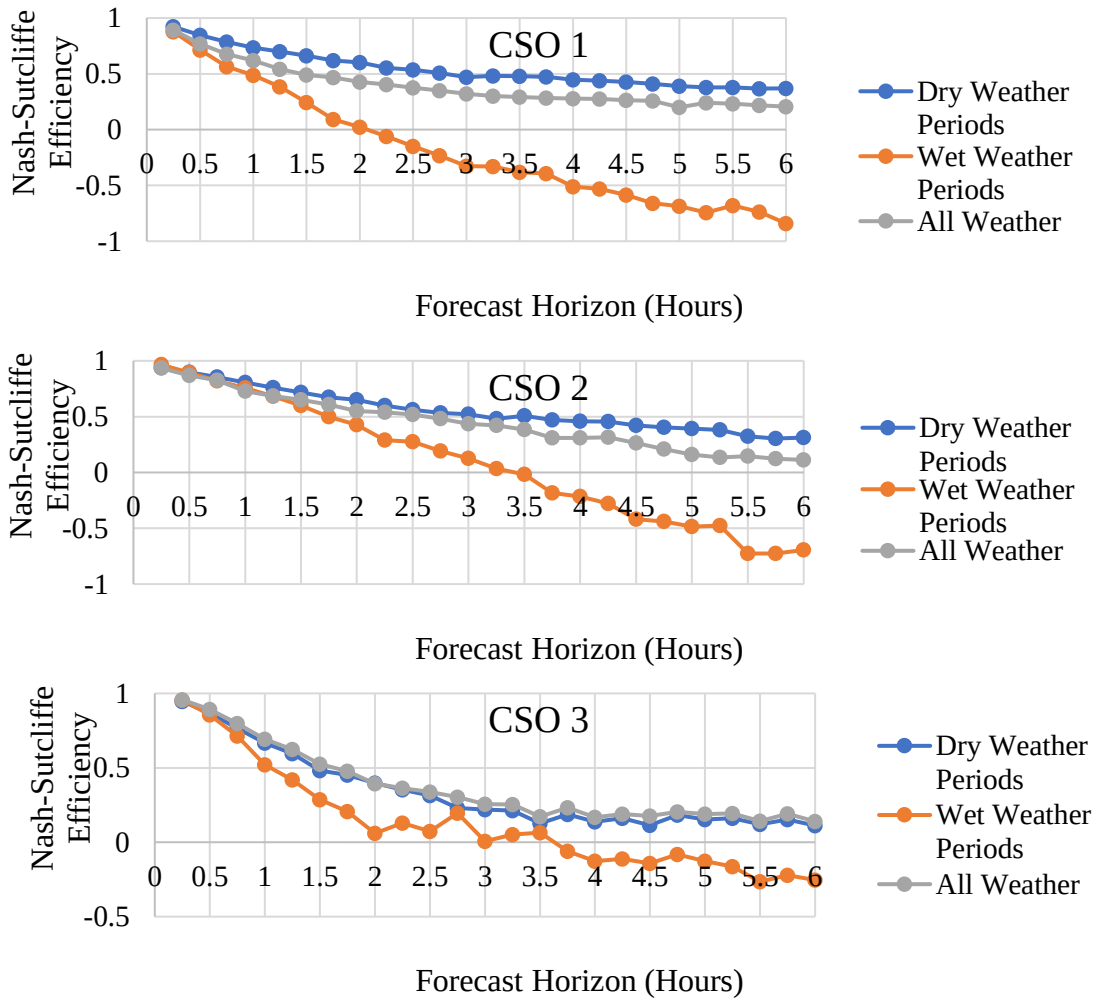


Figure 3-9 Comparison of ANN performance during wet weather timesteps, dry weather timesteps and all weather timesteps for three different CSO sites.

desire to use CSO level forecasting to understand CSO behaviour during heavy rainfall especially, so that they can anticipate spill events in advance and apply proactive measures, therefore this is a significant concern.

In recent years, the problem of learning from imbalanced data has drawn significant interest from academia and industry. A number of techniques are available to overcome data imbalance, for example, over or under sampling, penalised models and synthetic data generation (He & Ma, 2013). However, these approaches become more complex when working with time-series data as the time dependency among the observed values must be taken into account. A CEANN was utilised here as it can easily be applied to time series data. Additionally, it is important for the wastewater utility to forecast levels during extreme rainfall events as accurately as possible. Using a committee approach, an EANN can be specifically trained for wet weather, generating optimum results. Committee ANNs have been previously applied to hydrology and water resource-related problems, for example in the areas of groundwater salinity prediction (Barzegar & Asghari Moghaddam, 2016) and streamflow prediction (Lee & Kang, 2016) and have been demonstrated to be more robust and improve the generalization ability of a single ANN model.

The committee machine used here is a Bi-model CEANN. The first model is trained and tested using dry-weather data only (i.e. timesteps without rainfall) and the second is trained using wet-weather data only. The two models are therefore optimised specifically for dry weather and wet weather periods respectively, avoiding the problem of imbalanced training data. When forecasting in real time, the outputs from the two models are combined to give a single result, and the model is able to provide accurate CSO level forecast during both dry weather and during rainfall events.

It should be noted that committee ANNs generally combine more than two different models. Here only 2 networks are used as the main purpose of the committee is to effectively capture both dry and wet weather data. However, in the future additional networks could be easily incorporated if this was demonstrated to improve the results.

There is no universal definition for dry and wet weather periods. Dry weather is often defined using daily precipitation values, (e.g. in Polade et al., (2014) dry

weather is defined as under 1 mm for 24 hours). Others use weekly value; according to IWEM (1993) when the wastewater is mainly domestic in character, DWF is defined as seven consecutive days without rain following seven days during which the rainfall did not exceed 0.25 mm on any one day. Similarly Islam et al. (2017) use a definition of periods with under 5 mm of rainfall in a 7 day period. However, definitions over such long time periods will not work here. Instead a binary threshold has been used to categorise the wet and dry weather data based on cumulative rainfall over a past number of timesteps. The thresholding decision is defined as

$$Z_i = \sum_{t=0}^{n} R_{i-t}$$

$$\begin{cases} Wet & if\ Z_i > \theta \\ Dry & otherwise \end{cases}$$

where R_i is rainfall at time i, θ is the wet weather threshold and n is the number of past timesteps considered. θ and n are here set as 0.5 mm and 10 respectively. These values were determined by analysing historical CSO level and rainfall data, and manually identifying for which values rainfall during dry weather periods had a negligible effect on the CSO level.

There are a number of different methods available for combining the individual EANN outputs. The simplest and most common approach is simple averaging which assigns equal weights to all the forecasting component models (Lincoln & Skrzypek, 1989). Other methods include weighted averaging where the contribution of each input model is weighted according to its trust, confidence or estimated performance output (Jafari & Mashohor, 2011; Opitz & Shavlik, 1996), and majority voting where the correct result is the one chosen by the most neural networks (An & Meng, 2010).

In this methodology the aim of the model blending is to use the most appropriate model dependent on the rainfall at the current timestep. Therefore, a weighted averaging approach has been selected, dependant on the intensity of the rainfall of the current timestep. The models are combined using a non-linear weighted average based on the sigmoid function:

$$S_t = \frac{e^{x_t}}{1 + e^{x_t}}$$

$$L_t = S_t W_t + (1 - S_t) D_t$$

where *x* is the cumulative rainfall in the CSO catchment over a past number of timesteps, L_t is the overall ensemble output, W_t is the output from the wet weather model and D_t is the output from the dry weather model at time *t*.

A sigmoid function was selected due to its characteristic 'S'-shaped curve which exists continuously between 0 and 1, providing a continuous, or 'soft', transition between the two EANN models. During dry weather (i.e. when S = 0) only the dry weather model is utilised, during heavy rainfall (i.e. when S = 1) only the wet model is utilised, and during all other times a combination of the two models is used, weighted to the intensity of rainfall. The combination process is illustrated in Figure 3-10.

A discontinuous switching method was initially considered, whereby values for wet weather timesteps are obtained from the wet model only and for dry time steps the dry model only. Using this method, the threshold for designating timesteps as wet or dry is the same as that used to optimise the wet and dry EANN models, given above. However, unlike the CEANN, where the sigmoid function provides a continuous transition between the two EANNs this model uses

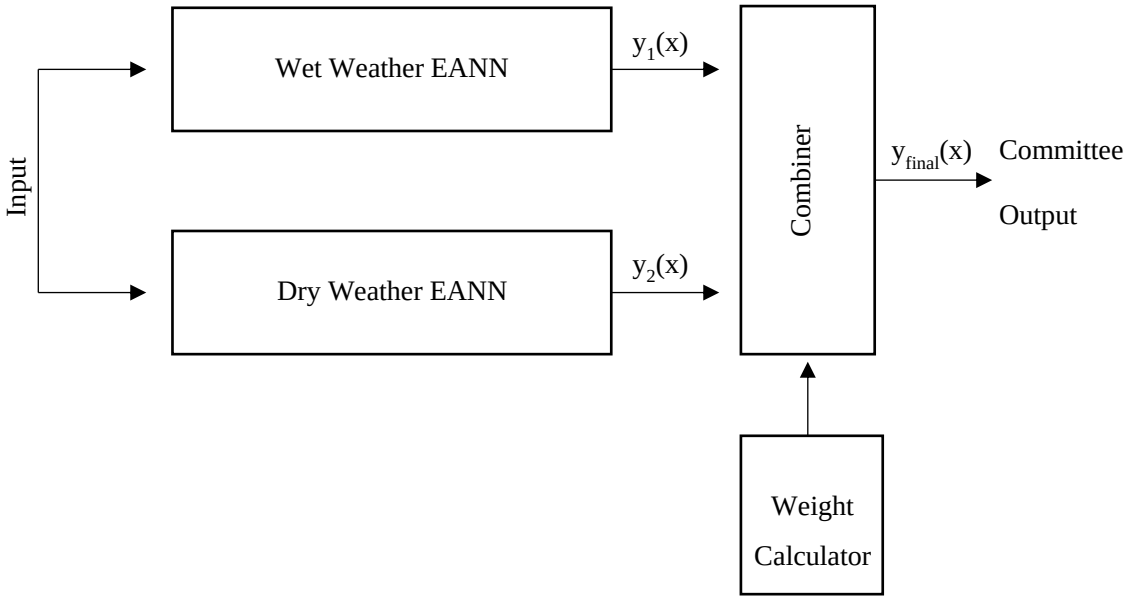


Figure 3-10 Schematic diagram of the CEANN model.

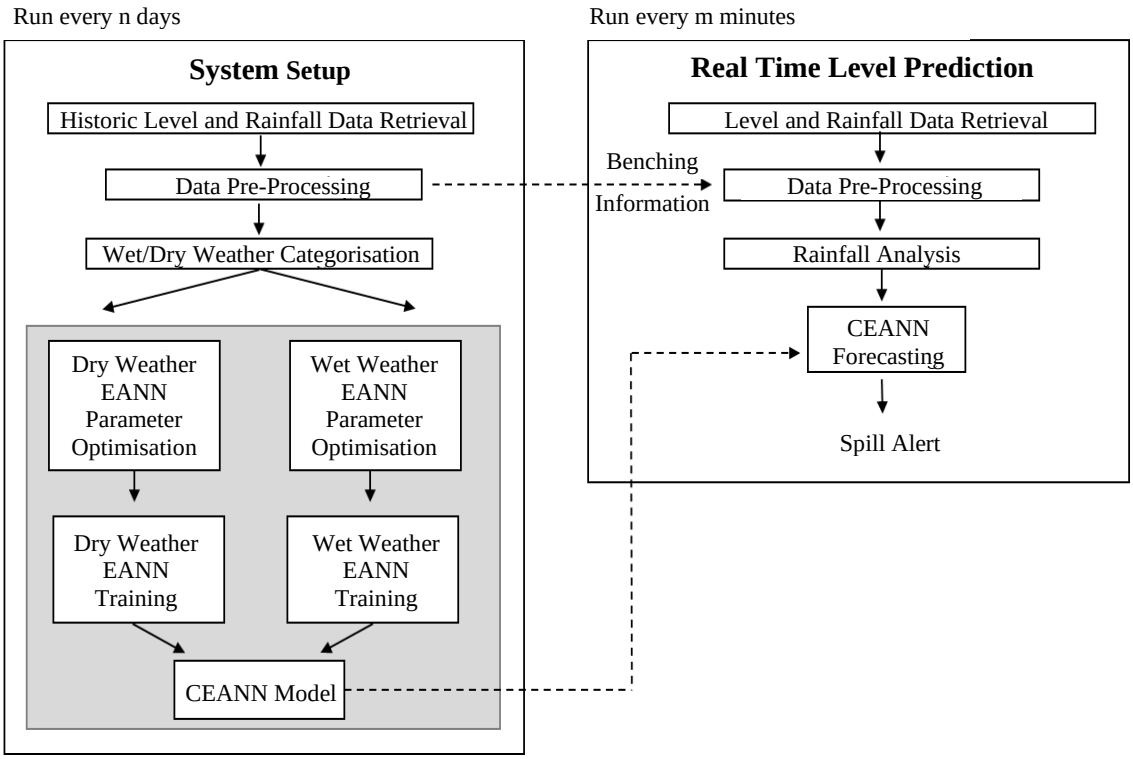


Figure 3-11 Diagrammatic representation of the CSO level prediction methodology.

a sharp transition, switching from one model to the other - thus it is possible that a small change in output from the two input models could produce a large change in the overall model output, which is undesirable. A representation of the CSO level prediction system, utilising the CEANN model is presented in Figure 3-11.

3.7 Summary

In conclusion, this chapter gives details of three ANN models developed to forecast sewer levels in a CSO chamber;

(i) A simple ANN constructed using the trial and error method for determining network structure and parameters

(ii) An EANN model optimised for all weather types

(iii) A bi-model CEANN composed of two evolutionary artificial neural network (EANN) models, optimised for wet and dry weather respectively, and combined using a non-linear weighted averaging approach.

A novel methodology for automatically removing erroneous level data due to benching has also been developed. Benching can cause CSOs to mistakenly appear to be overflowing. This methodology will therefore be of value to wastewater utilities who have a duty to minimise the number of overflows in their network and are required to report the frequency and duration of CSO spills.

The prediction models are designed to operate quickly and efficiently, making use of rainfall and CSO level data which are typically routinely collected by wastewater utilities. The models have the potential to be used beneficially by wastewater utilities to model CSO levels in the wastewater network in near real-time and provide alerts for upcoming spills. This will enable better decision making and proactive management of overflow events.

The different models are applied to case study sites in Chapter 6 and their capabilities are presented and compared. Additionally, the results of sensitivity data analyses that focussed on evaluating the performance of the methodologies presented in this chapter for different choices of parameters and methods are described.

The CSO level forecasts are also utilised in the blockage detection system methodology, described in the following chapter.

Chapter 4: Blockage Detection Methodology

4.1 Introduction

Sewer blockages are a major issue in cities around the world. Blockages cause numerous issues such as unplanned maintenance, internal and external flooding, significant pollution, costly clean-up and compensation costs and threats to public health due to water borne pathogens (Korving et al., 2006; Mcintyre et al., 2012). A number of studies have indicated that blockages are responsible for a significant proportion of the failures which occur in wastewater networks (Jin & Mukherjee, 2010; Xie et al., 2017). In the UK alone there are approximately 300,000 blockages every year, resulting in costs of £100 million. Ofwat imposes a statuary duty on wastewater utilities in England and Wales to minimise the frequency of blockage events. However, with deteriorating sewer networks and increased water efficiency the number of sewer blockages suffered on public sewer networks is increasing. The timely prediction of blockage events therefore plays an important role in the management of urban water systems.

As part of the Event Duration Monitoring (EDM) program implemented by the Environment Agency wastewater utilities have been required to monitor levels at the majority of CSOs in England and Wales since 2020. As a result, large quantities of increasingly accurate level sensors have been installed in the sewer network. In conjunction with reductions in data storage costs and improved computer processing power this has led to water utilities routinely collecting large volumes of accurate sewer level data in near real time. This data is extremely valuable, however it is not feasible for human operators to process this raw data manually. This motivates the demand for intelligent data-driven analysis to assist in real-time data-driven modelling and management of the wastewater system.

Historically, wastewater utilities have relied on their customers to report blockage events, employing a reactive repair and maintenance approach. However, this results in increased loss of service and customer complaints, in turn affecting

regulatory performance. Additionally, if the blockage does not cause any external effects (e.g. flooding) it can remain undetected for many months, increasing the likelihood of combined sewer overflows into nearby watercourses. There are also a number of hardware-based techniques available to detect blockage events, such as CCTV, acoustic techniques and laser profiling. However, these techniques are, generally, expensive, intrusive and time consuming to implement and require the added expensive of a human operator.

Much of the recent published on sewer blockage management has focused on predicting blockage rates and representing the spatial distribution of sewer failures and/or flood hazards (Caradot et al. 2011; Cherqui et al. 2015; Post et al. 2017) using statistical methods (Pulido et al., 2019). These methods are extremely valuable in improving planning and scheduling of asset inspection, maintenance, and replacement programs in sewer systems so that blockages are less likely to form. However, they are not capable of detecting blockage events in real time. There is a need for technologies able to monitor the sewer system in real time and detect blockages as soon as they occur, allowing utilities to proactively remove obstructions before the customer and environment are affected.

There are some recent commercially available products designed for blockage detection. The company Detectronic offer a service to predict blockage events and provide early warnings of network failure by monitoring sewer data in real time using human data analysts (Detectronic, 2018, 2019). The system develops normalised data profiles for all monitored sites and develops predictive criteria for early preventative intervention. The systems use wastewater flow meters in conjunction with level monitors. Wastewater utilities are required to install level monitors under the EDM program, however, flow monitors are not required, and are more expensive than level monitors. Utilities would therefore require the additional expense of installing flow monitors to implement the Detectronic system. SMART Sewer™ developed by Environmental Monitoring Solutions UK detects developing sewer blockages autonomously in real-time using level data (EMS, 2020). The system utilises fuzzy logic and is designed to be deployed in high incidence and high-risk areas, enabling the optimisation of sewer cleaning and blockage removal. As both these systems are commercially available products there is little literature available describing the methodology of the

systems in more detail or showing their performance. They have been applied to case studies by wastewater utilities, for example Southern Water trialled the SMART sewer in hotspot areas in Brighton and East Worthing (Southern Water, 2018), but such new technology has not yet become an industry standard. They do, however, demonstrate that utilities see the potential in such systems and are willing to use them in the future.

A novel methodology is presented here which is designed to detect blockages and other unusual events in the proximity of CSOs, in near real-time. The proposed event detection system (EDS) applies two different detection techniques; EANN discrepancy analysis and statistical trend-based analysis and is designed to provide automated real-time alerts for blockage events. The system is designed to make use of the loggers already installed as part of the EDM project and does not required the installation of any additional sewer system monitors. It is envisioned that the methodology will be beneficial to water and wastewater utilities, allowing proactive management of blockages before they affect the customer and the environment.

4.2 Detection System Overview

Currently wastewater utilities employ human operators to determine if blockage events have occurred in the sewer network. Generally, these operators are only alerted to the presence of a blockage event when internal or external flooding or a combined sewer overflow occurs. The event detection system presented here is an expert system designed to mimic the behaviour of a qualified and experienced human operator, monitoring sewer level data and detecting blockage

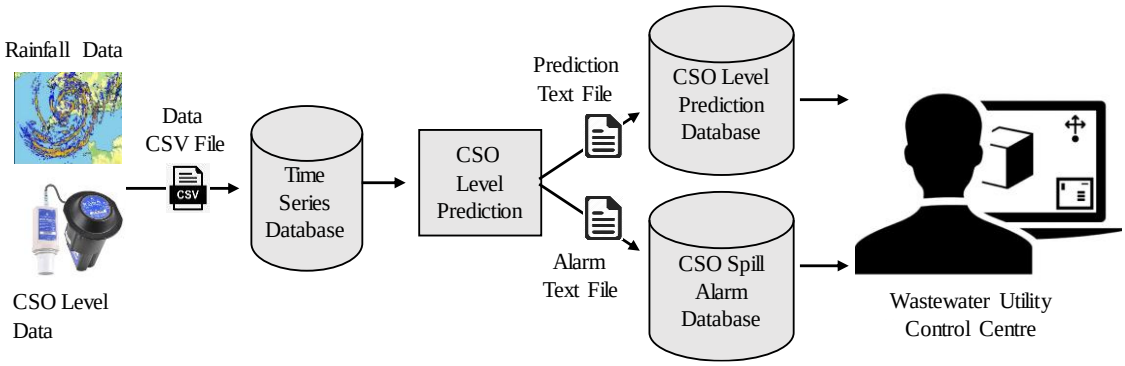


Figure 4-1 Schematic representation of the online implementation of the CSO level prediction methodology.

events a timely and reliable manner. Expert systems such as this offer a number of benefits when compared with human experts. They are inexpensive to operate, easy to reproduce and distribute, and can provide permanent documentation of the decision process. Expert systems can also produce consistent results for the same tasks and handle similar situations consistently.

Figure 4-1 provides a schematic of the online implementation of the event detection system. For each CSO chamber site, sewer level data is collected from the level sensor and rainfall data is retrieved from the Met office at regular time intervals. The data is stored in a set of automatically updated comma-separated values (CSV) files in a time-series database. At each timestep the event detection methodology is performed, as described below. If a blockage event is determined to have occurred, an alarm is generated, and the user notified.

The detection system operates by identifying abnormalities in CSO level data in real time. During normal functioning of the wastewater network sewer levels display predictable behaviour – during dry weather the levels exhibit a steady diurnal pattern and during rainfall events the sewer levels increase, resulting in

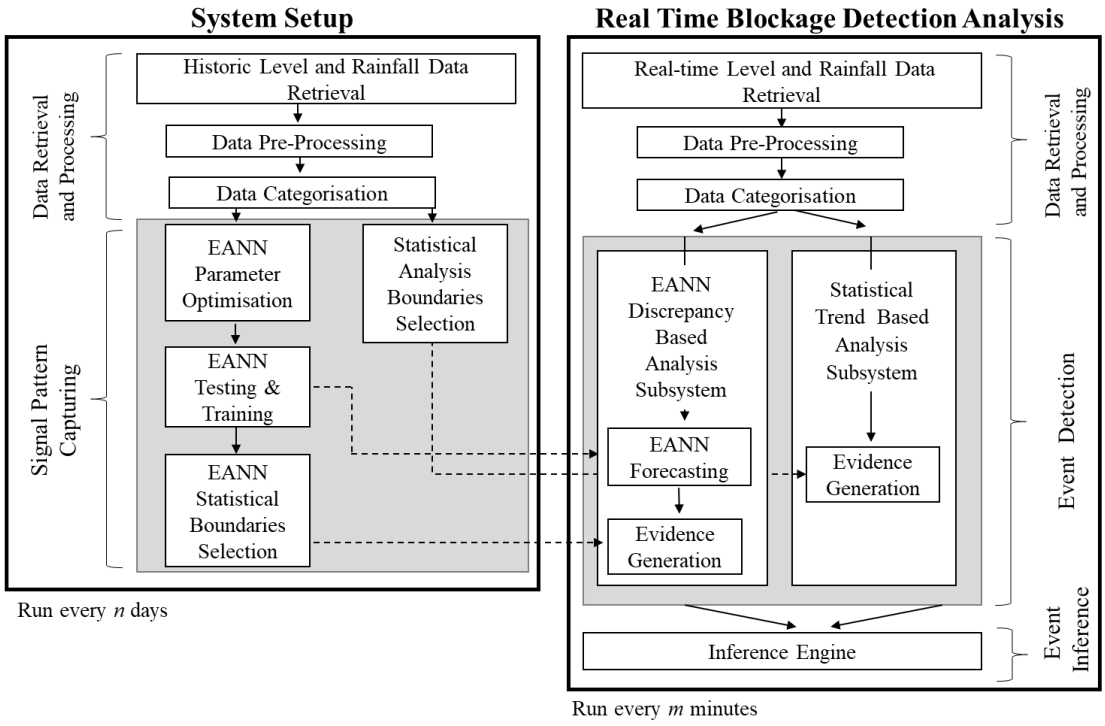


Figure 4-2 Diagrammatic representation of the blockage detection methodology.

an overflow if the level in the CSO exceeds the spill height of the chamber. However, when a blockage, or another abnormal event, occurs in the proximity of a CSO chamber this can cause abnormalities in the level behaviour, resulting in an increase in level if the blockage is downstream of the chamber and a decrease if the blockage is located upstream. Therefore, by continually monitoring the actual CSO level and identifying behaviour which deviates from what is expected, blockages can be detected in real time.

Figure 4-2 provides a visual representation of the EDS methodology. The framework of the blockage detection system consists of three subsystems: (1) data retrieval and processing, (2) event detection, and (3) inference. The first subsystem consists of retrieval and pre-processing of the incoming data for the CSO chamber being analysed and categorisation of the data. The second subsystem is designed to generate evidence that a blockage has occurred. Two separate methodologies are employed to analyse the level in the CSO chamber. An EANN discrepancy-based analysis module is designed to monitor the difference between EANN level predictions and the observed CSO level and a statistical trend-based analysis module uses statistical analysis to monitor trends in the CSO level data. The aim of using a combination of statistical and AI techniques is to provide more rigorous evidence that a blockage has occurred. Different types of blockage events have different effects on sewer levels. For example, sedimentation and build-ups of fats, oils and grease (FOG) cause gradual changes in level, whilst sewer collapse and snagging of objects generally cause sudden fluctuations. The application of different types of detection techniques increases the likelihood that all types of blockage events are detected quickly and reliably.

The third and final subsystem consists of an inference engine. The engine analyses the information from the two evidence generation modules to determine if there is sufficient evidence to determine that a blockage event has occurred, and thus generate a blockage alert and notify the user. Additional information concerning the blockage is also provided by the system to facilitate the user in deciding how to respond to the event. Each of the three subsystems will be discussed in detail in the following sections.

The detection system has two modes of operation: a set-up (training, offline) mode and real-time (deployment, online) mode. The set-up mode is devised to enable the system to 'learn' the behaviour of the CSO being analysed. Here the system uses historical data to capture normal CSO levels. It is advised that this mode is applied periodically and also following significant sewer system configuration changes. This automatic retraining is necessary as the behaviour of the wastewater system may change over time. For example due to planned changes to the sewer network and wastewater treatment plants or the construction of new buildings connected to the network. This is a problem known as concept drift (Widmer & Kubat, 1996). The EDS system must therefore be able to adapt to this new definition of 'normal' in an unsupervised, automated fashion.

It is important to note that the system should not be retrained to adapt to changes caused by undesirable and unintended modifications to the network - for example a change in sewer level caused by a build of siltation in the pipes which has not been identified and removed by the wastewater utility. The system will then incorrectly view this new 'abnormal' behaviour as normal, impeding its ability to detect future blockage events. This is related to the problem of catastrophic forgetting – the tendency of ANNs to completely and abruptly forget previously learned information upon learning new information. It is anticipated, however, that this should not be a big concern here - any such measurable changes to sewer behaviour should be detected by the EDS and so wastewater utilities will be aware of their existence. The wastewater utility can therefore ensure that the system is not retrained using this abnormal data. There are also methods designed to overcome catastrophic forgetting, such as progressive neural networks and elastic weight consolidation. A good summary is presented by Chen & Liu, (2018). These techniques have not been applied here, however they could be included in the future.

The second mode of operation is the real-time mode. This is the normal online operating mode used to detect events and raise alarms in real-time. The system is designed to operate at the communication interval of the real-time level and rainfall data received by the system, e.g. operate at 15-minute intervals for 15-minute rainfall data.

The process for constructing the EDS involves the selection of several parameters. The choice of these parameters controls the performance of the detection system. The following main factors have been analysed when investigating the parameter choice: the true positive rate (i.e. sensitivity) of the system, here defined as the proportion of blockage events correctly identified, the false positive rate of the system, here defined as any blockage alarms generated when there is no blockage present, the false negative rate, defined as any blockage events which are not detected, and the event detection time, defined as the time difference between when a blockage event first occurs and when an alarm is first raised.

The true positive and false positive rate are calculated as

$$True\ Positive\ Rate = \ Sensitivity = \frac{number\ of\ positive\ instances\ correctly\ classified}{total\ number\ of\ positive\ instances}$$

$$False\ Positive\ Rate = \ 1 - Specificity \ = \frac{number\ of\ negative\ instances\ misclassified}{total\ number\ of\ negative\ instances}$$

where the specificity (i.e. the true negative rate) of the system measures the proportion of negatives that are correctly identified.

In general, there exists a trade-off between minimising the detection time and the false negative rate of the EDS, whilst also minimising the number of false alarms generated. The sensitivity of the system must therefore be carefully selected; an overly sensitive system results in wasting operational resources on investigating false alarms, whilst an insensitive system causes excessive losses due to missed events and long detection delays.

Finding an optimal threshold is a challenging problem and has been discussed extensively in the literature. Discussions with various different wastewater industry personnel were used to inform the parameter selection process here. Overwhelmingly the industry personnel emphasised that it is often difficult to encourage workers to use new technology. This is especially true if the technology generates false alarm and is viewed to be ‘faulty’ or ‘untrustworthy’ and causes unnecessary work. This may result in users losing confidence in the system. Reducing the number of false alarms is also very important when using a system in real time as it reduces the time spent on manual quality control.

When selecting the system's parameters it was therefore decided to emphasise minimising the false alarm rate significantly, whilst still maintaining a good rate of detected events. The few blockages missed due to implementing high evidence thresholds are likely to be small and have a less significant impact on the CSO level, and therefore are unlikely to cause overflow events or flooding.

Currently the EDS methodology has only been applied to downstream blockages, which cause an increase in CSO level. This is because an analysis of CSO level data found very few upstream blockages. Indeed, upstream blockages made up only 7.6% of the total blockages identified in the analysed wastewater network. It is possible that either blockages are less likely to form upstream of the CSO chamber, or that upstream blockages do not have a significant effect on CSO level and therefore cannot be detected in the level data. Regardless, the small number of upstream blockages means there is not enough data available to reliably select system parameters for these events. It is possible that in the future, when more level data is available, parameters can be selected for upstream blockage using the same approach described here. The point, however, is that the methodology shown here can be used for the detection of blockages forming upstream of CSO chambers as well.

4.3 Data Pre-processing and Categorisation

The first subsystem of the EDS detection process consists of data pre-processing and categorisation. Data processing is performed to construct a clean, uniform time series, i.e. with any anomalies identified and removed and with regularly spaced timesteps. It is especially important that any erroneous CSO level data is removed, as it could be mistaken by the EDS as abnormal behaviour caused by a blockage event and therefore trigger a false alarm. The data pre-processing is conducted here as described in Section 3.2.

Once the data pre-processing is performed the data is then categorised. Each timestep is assigned a category based on its corresponding rainfall data characteristics. Water levels in the sewer are highly dependent on the characteristics of the rainfall in the surrounding catchment and different rainfall events cause different sewer level dynamics. Blockage detection is performed by analysing levels in the CSO and distinguishing unusual behaviour caused by

blockage events from normal, every-day behaviour. By categorising the data according to rainfall type, the blockage detection methodology can be tailored for the different types of rainfall events which occur, thus increasing the sensitivity of the system, minimising the event detection time and decreasing the number of false alarms.

The rainfall categorisation first requires the identification of distinct rainfall events. Separating continuously recorded rainfall data into independent events can be a difficult task, and often relies on subjective interpretation. It is necessary to establish clear methods to determine the beginning and end of each rainfall event. A poor identification of events may lead to invalid statistical results, thus leading to flawed analyses. We have applied the inter-event time definition (IETD) here to identify rainfall events. The IETD (also known as the inter-arrival time or inter-storm period) is arguably the most commonly used approach. It refers to the minimum dry period that can adequately divide two rainfall events (Joo et al., 2014). Analysis into the use of the IETD has been conducted since the 19th century, and this method has been frequently applied to urban drainage systems (e.g. James 1994; Lee and Kim 2018). An example of the identification of rainfall events using the IETD method is illustrated in Figure 4-3. Rainfall groups A and B are considered separate events as the time interval between them is greater

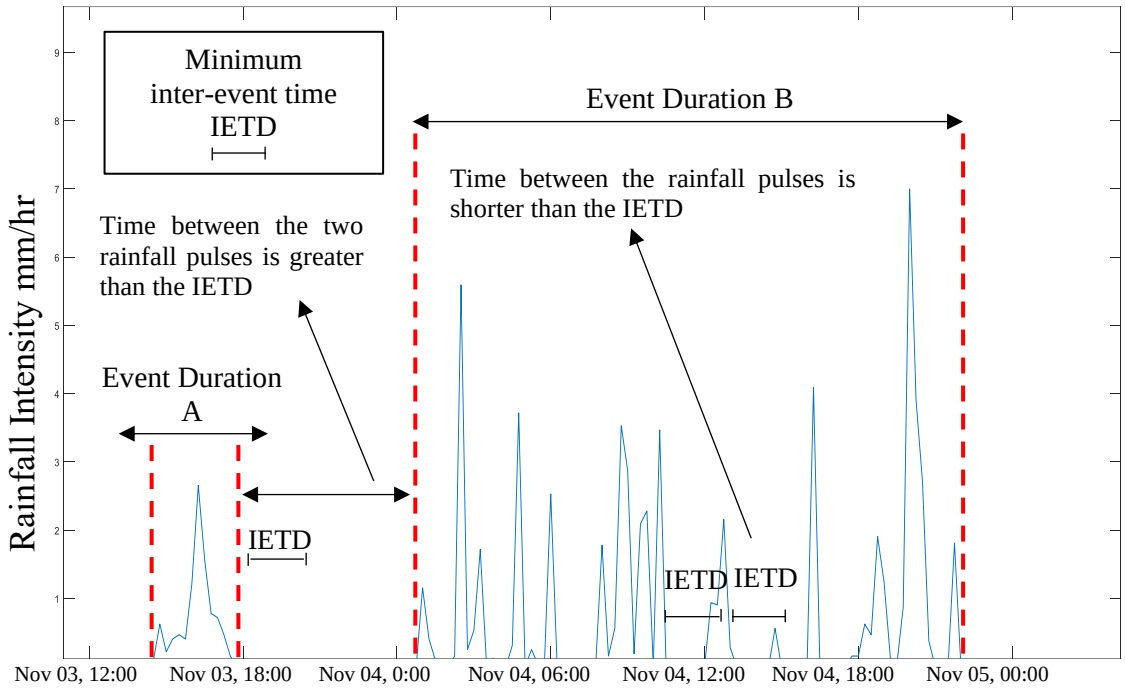


Figure 4-3 Illustration of identification of rainfall events based on the IETD.

than the IETD. Although rainfall event B contains many individual pulses of rainfall they are regarded as a single event, as the interval time between the separate pules is shorter than the IETD.

Once the rainfall events have been identified using the IETD criterion they can then be categorised. The characterisation and categorisation of rainfall is important in many fields, including flood forecasting (Alfieri et al., 2015), urban drainage (Balbastre-Soldevila et al., 2019; Chen & Adams, 2005) and infrastructure design for buildings and roads (Cheng & Aghakouchak, 2014). Rainfall varies considerably from one country to another, and as such no universal rainfall classifications exist. However, rainfall events are most commonly characterised using rainfall Intensity-Duration-Frequency (IDF) curves (Koutsoyiannis et al., 1998). The data is categorised here according to the rainfall duration and intensity only. This is because the rainfall categorisation is performed to construct a profile of normal sewer behaviour for the different rainfall classes. Dividing the data into too many rainfall classes could result in categories with insufficient data to construct an adequate profile, whilst dividing into too few classes would result in ineffective detection methodology (we tried not using rainfall classes at all initially, but this did not work out well due to different nature of sewer and CSO responses to different rainfall events). It was decided therefore to not use frequency (i.e. return period) as a factor as it resulted in dividing the data into too many categories.

An analysis was performed to determine the number of categories which produced the best results when utilised by the EDS, presented in detail in Section 4.3.1. Based on this analysis the rainfall is divided in this work into 6 principal categories, described as follows:

(i) Dry weather (i.e. no rainfall),
(ii) Short duration, low intensity rainfall (i.e. type A),
(iii) Short duration, medium to high intensity rainfall (type B),
(iv) Long duration, low intensity rainfall (type C),
(v) Long duration, medium to high intensity rainfall (type D) and
(vi) Post event period (i.e. the period immediately following a rainfall event).

The categories are illustrated in Figure 4-4. The post event period is designed to capture the period immediately following a rainfall event when the water level in

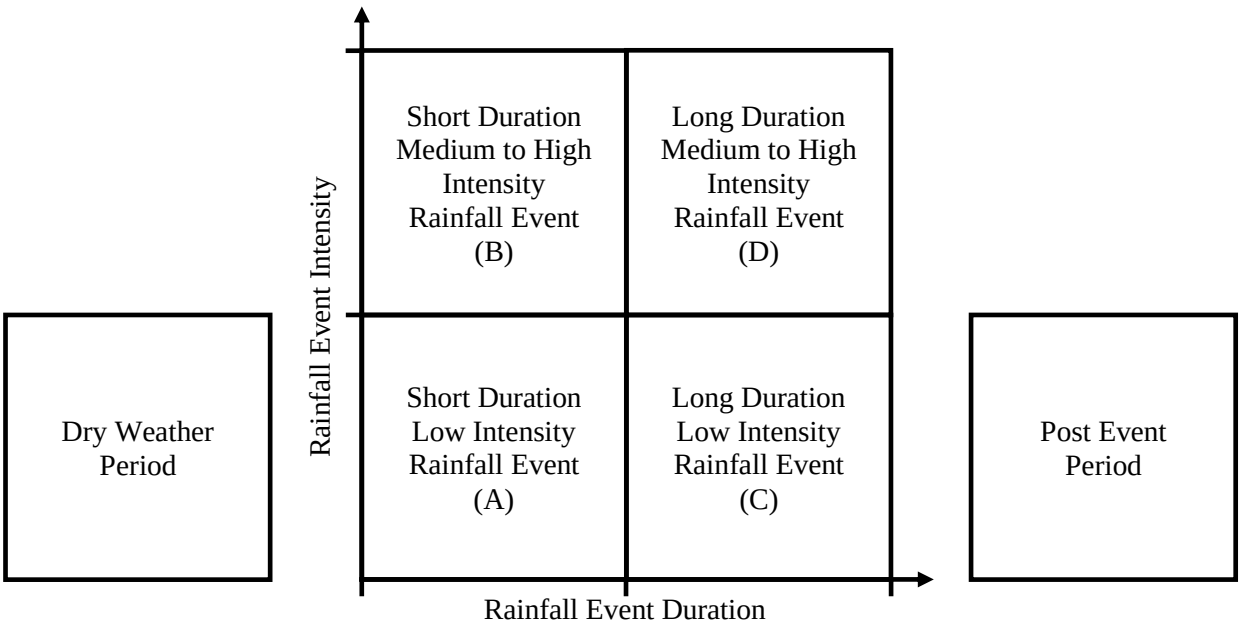


Figure 4-4 Matrix for determination of the rainfall event classification.

the sewer remains high. This is due to the time it takes for water to flow from the surrounding catchment to the CSO chamber.

Further details, i.e. the actual rainfall intensity and duration thresholds used here to form above classes are given in the following section.

4.3.1 Rainfall Categorisation Parameter Selection

This section presents the process for the selection of the duration and intensity thresholds selected for the 6 rainfall categories and the selection of the inter event time definition. As stated above, due to the disparity in worldwide precipitation no universal rainfall classifications exist. Different thresholds for rainfall intensity classification have been created for use in studies for various locations. e.g. Italy (Caloiero et al., 2016), the Mediterranean (Alpert et al., 2002), Korea (Sohn et al., 2013) and China (Yuan et al., 2015).

Figure 4-5 displays an example of the typical distribution of rainfall duration and intensity, for a site located in North-West England. Figure 4-5(a) presents a frequency histogram of 9 years of rainfall intensity data. The data was obtained from the UK MET office, measured in mm/hr and was collected at 5-minute intervals. The frequency refers to the distribution of the rainfall intensity datapoints. As can be seen the majority of the rainfall is low intensity (under 0.5 mm/hr). Figure 4-5(b) presents a frequency histogram of the distribution of the

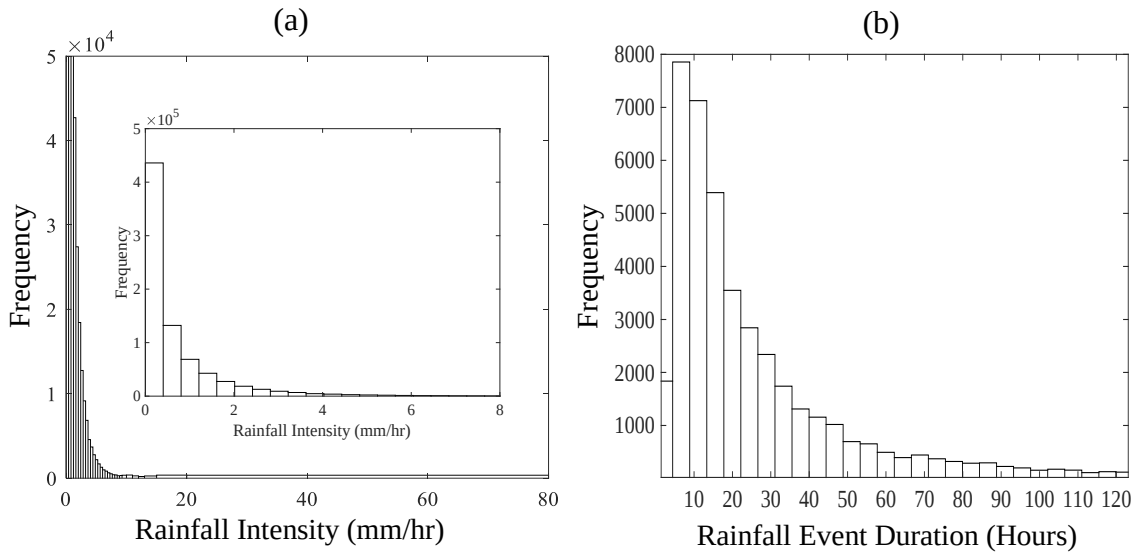


Figure 4-5 (a) Histogram of rainfall intensity and (b) histogram of rainfall event duration, from a CSO site.

duration of the rainfall events identified in the same 9-year dataset. The majority of events are under 8 hours and the most common event durations are between 2 and 4 hours.

The UK Met Office (Met Office, 2007) classifies rainfall (excluding showers) as ‘slight’, ‘moderate’, or ‘heavy’, for rates of accumulation less the 0.5 mm/h, 0.5 to 4 mm/h, and greater than 4 mm/h, respectively. Showers, which are characterized by short duration and rapid fluctuations of intensity, are classified as ‘slight’, ‘moderate’, ‘heavy’, or ‘violent’ for rates of accumulation of 0 to 2 mm/h, 2 to 10 mm/h, 10 to 50 mm/h, and greater than 50 mm/h, respectively.

There are different approaches for creating precipitation intensity classifications (Yuan et al., 2015). The simplest is to use ‘ad hoc’ threshold values, e.g. as used in Qian et al. (2007). However, the threshold values chosen this way are rather subjective in nature and hence hard to justify. Another approach uses percentiles of precipitation distribution; low percentiles of precipitation are defined as light precipitation and high percentiles as heavy precipitation (Karl & Knight, 1998). This approach is useful when analysing precipitation extremes, but again suffers a similar problem of subjectivity in defining the threshold values. A third approach is to categorise the precipitation intensity into several bins. For example Alpert et al. (2002) divided daily precipitation into six bins by values in powers of 2, whilst

Osborn et al. (2000) divided precipitation into 10 levels with each level contributing to 10% to the total precipitation.

It was decided here to use a trial and error approach based on five criteria identified to be important for the methodology. The five criteria are as follows:

i. Distinctiveness: Each rainfall category must produce a distinct CSO level profile. This avoids proliferation of redundant categories and encourages a compact model.
ii. Coverage: The categories must encompass the full range of rainfall conditions.
iii. Sufficient data: Each category must contain sufficient data samples to construct an adequate profile of the daily CSO level when applied to a minimal 4-month dataset (the minimum length of data recommended for training the EDS for a particular CSO). This is most important during high duration, moderate to high intensity rainfall events (i.e. rainfall type D) which are considerably less frequent than other rainfall types. High intensity rainfall events also cause more changeable and unpredictable sewer behaviour. A poor understanding of the sewer levels during these types of rainfall could result in a large number of false alarms during real time running of the system.
iv. Time limited: When the system is run in real-time, forecast rainfall data is used to determine the rainfall category of the current timestep based on the predicted rainfall. Therefore, the threshold between short and long duration rainfall cannot be greater than the longest forecast horizon. The data used here is 6 hour ahead Met Office forecast rainfall data – therefore the event duration threshold may not be greater than 6 hours.
v. Semantic meaning: The categories should be human-readable. This is important when explaining the system and its results to a user. This criteria aligns with the trend for explainable AI (Wojciech et al., 2019). Semantically meaningful categories are also helpful to any future developer, who may maintain, adapt, or extend the system.

The intensity and duration thresholds were selected using these criteria, based on an analysis of rainfall from 30 different CSO sites located in North West England. Various different sets of parameter values were considered, from 0.5 to

10 mm/hr at an interval of 0.25 for rainfall intensity and from 1 to 5 hours at 1 hour intervals for the event duration and evaluated using a grid search method to identify the best combination across all the CSO sites. The threshold for moderate to high intensity rainfall was thus set at 2 mm/hr and the threshold for long duration rainfall set as 3 hours. These thresholds appear to correspond well with the rainfall distribution presented in Figure 4-5.

The sufficient data criteria was found to be an important consideration as varying the parameters affected the numbers of events assigned to each category significantly. Indeed, the combination of the highest rainfall intensity and longest event duration thresholds analysed (10 mm/hr and 5 hours) resulted in zero events classified as type D rainfall for a number of sites. The selected parameters ensure that there should reliably be sufficient type D events to construct an adequate CSO level profile.

The selected rainfall intensity threshold may seem low compared to other established thresholds used for urban drainage. For example the minimum rainfall intensity used as a basis for design for buildings according to the British Standard for drainage design of gravity systems inside buildings (BSI, 2000) is an order of magnitude higher than the intensity threshold selected here. However these standards are set to ensure that the drainage system can accommodate rare extreme rainfall events which could otherwise result in serious flooding or damage. The threshold selected in this thesis is utilised only to facilitate making the system more adaptive to varying rainfall conditions. Ensuring that there is enough data in the moderate to high intensity rainfall classes to adequately construct bounds of normal behaviour is more important than accommodating these very uncommon rainfall events. Additionally, these standards generally consider short periods of rainfall (for example BSI (2000) gives values of design rainfall intensity for storm events lasting 2 minutes), whereas the methodology utilised here considers the average rainfall intensity of rainfall events over their whole lifetime, which is generally a few hours.

Regarding the selection of the inter-event duration value – again no general method is available. A number of different values have been suggested in the literature, from 15 minutes (Carbone et al., 2014) to 24 hours (Gaál et al., 2014).

According to Chin et al. (2016) the determination is highly dependent on the purpose and objectives of a particular study.

The IETD was set here as 2 hours, based on an analysis of the cross correlation between rainfall and level data from the 30 CSO sites. Cross correlation is a time-series analysis technique used to measure the similarity between two signals in relation to their time lag. Cross correlation has been used previously to investigate the response of CSO levels to rainfall (Fernando et al., 2006; Mounce et al., 2014a) and to analyse the response of the water-table in general (Lee et al., 2006; Mackay et al., 2014).

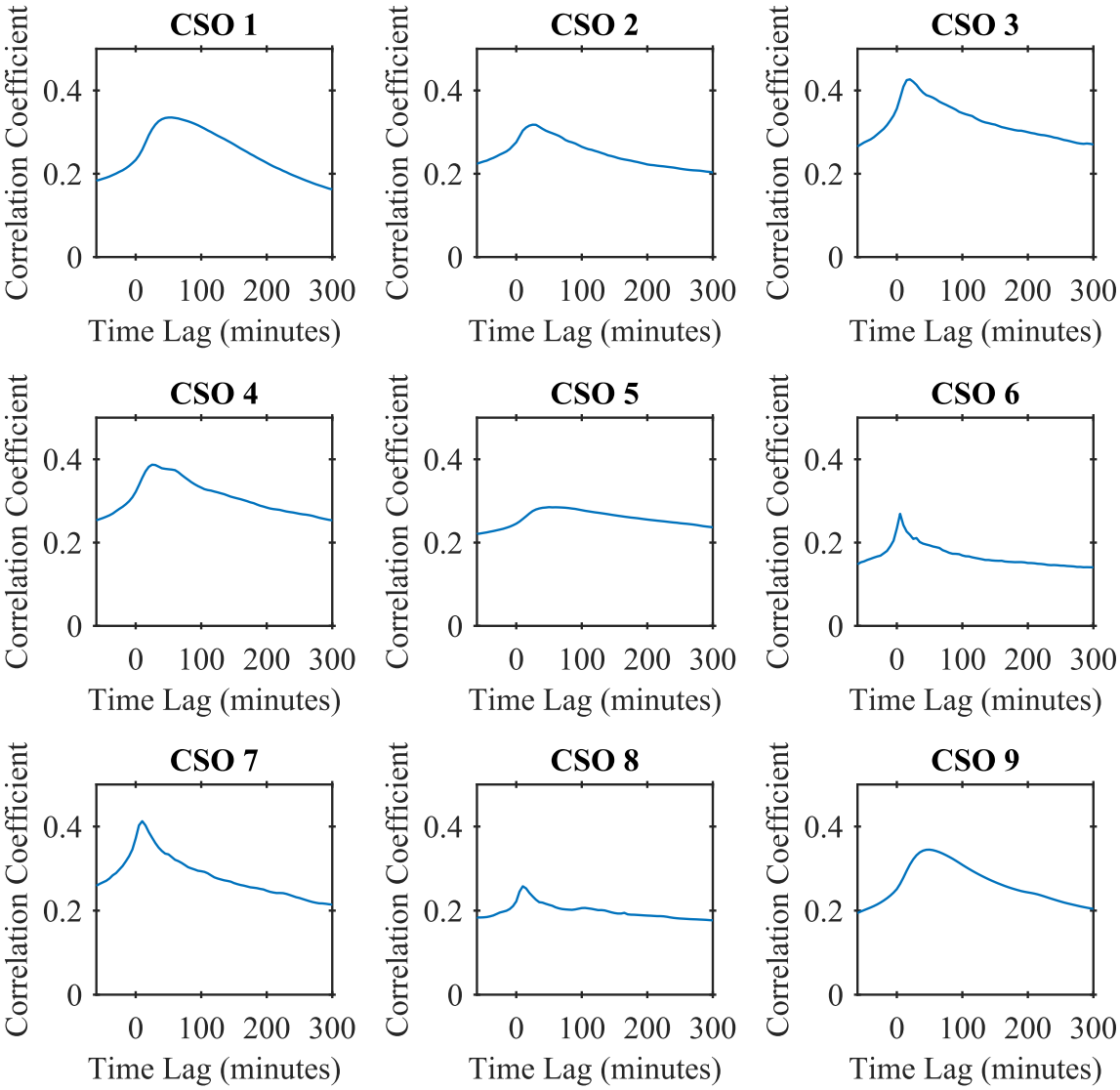


Figure 4-6 Cross correlation between CSO level and rainfall data for different CSO sites.

The cross correlation is calculated using the following relation (Davis, 2002):

$$C_t = \frac{n\sum RL - \sum R \sum L}{\sqrt{[n(\sum R)^2 - (\sum R)^2]\,[n(\sum L^2) - (\sum L)^2]}}$$

where C is the cross correlation at time lag t, n is the number of overlapping data points, R is the rainfall data and L is the CSO level data.

The cross correlation was analysed to provide information on the time taken for rainfall to reach the CSO chamber for different CSO sites. This is related to the time of concentration of the catchment – the time required for runoff to travel from the hydraulically most distant point in the watershed to the outlet. An example of the correlation produced for eight different sites located in northern England is displayed in Figure 4-6. The results provide a good indication of different CSO level responses to rainfall events.

This information was also used to select the value for the post event period duration – as the post event period is influenced by the time it takes for rainfall from the surrounding catchment to flow to the CSO chamber. The post event period was thus set as 3 hours. This value, in combination with an inter-event duration of 2 hours, appeared to produce generally good results for all CSO sites.

Selecting a unique IETD value for each CSO site was initially considered. According to Joo et al. (2014) when determining the IETD of an urban catchment the time of concentration, which varies from catchment to catchment, should be taken into account. Joo et al. define the IETD as the time period from the end of the rainfall event to the end of the direct runoff. There are many methods available to calculate the time of concentration, but they generally require physical information regarding the catchment hydrology which is not always available. Additionally, the blockage detection system is designed to operate using level and rainfall data only. However, as stated above the cross correlation can be used to estimate the time of concentration of a catchment and requires only level and rainfall data. Thus calculating a unique IETD value based on this information for each CSO site would be possible and could be performed in an automated fashion as required.

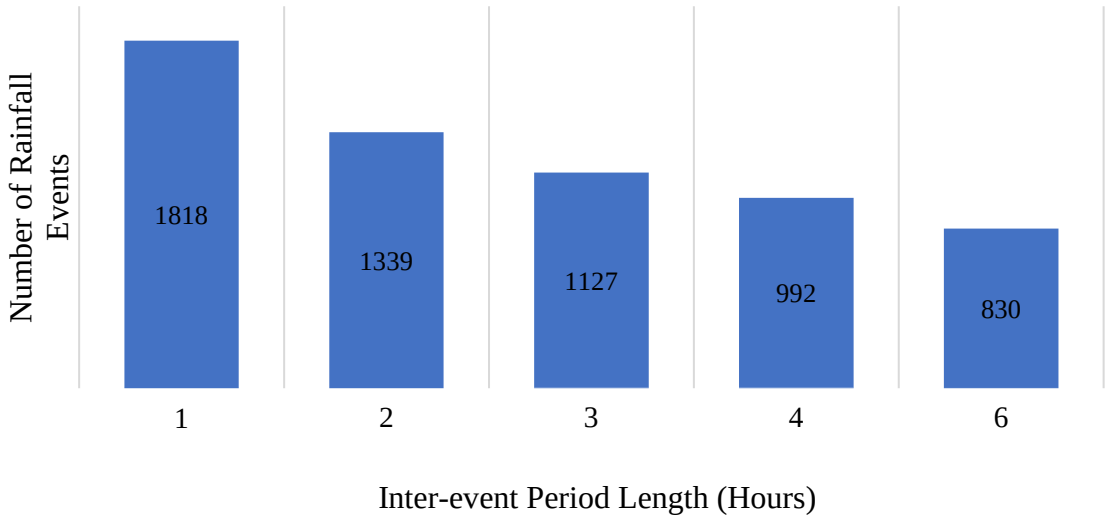


Figure 4-7 Effect of varying the inter-event period length on the number of rainfall events identified in a rainfall dataset.

Ultimately, however, it was decided that selecting a generic IETD value was preferable. Firstly, this approach is simpler, and secondly, selecting a unique value for each site the methodology is applied to could negatively affect the selection of generic parameters in other parts of the detection system. Varying the value of the IETD measurably changes the statistics of the identified rainfall event characteristics, such as rainfall volume, intensity and duration (Adams et al., 1986). An example of this is shown in Figure 4-7 which demonstrates how varying the duration of the IETD from 1 to 5 hours effects significantly the number the rainfall events in a nine-year dataset.

An analysis of a number of CSOs across the analysed wastewater network in both rural and urban locations demonstrated that these generic rainfall categorisation parameters can be applied to sites with varying characteristics and produce generally good results. It is likely that the values are also acceptable for the whole of the UK, although this has not been tested.

4.3.2 Rainfall Categorisation Process

The categorisation of historical rainfall data is a rather straightforward process once suitable threshold values have been defined. Real time categorisation is more complex – the rainfall events are categorised according to the duration of the event and the average rainfall intensity of the event. However, when the system is running in real time the eventual duration and intensity of an ongoing

event is not yet known and so it is impossible to use this information to categorise the event. Indeed, for time-steps with a rainfall intensity of 0 mm/hr where the time elapsed since the last period of rainfall is shorter than the IETD, it is not known whether the previous event is still ongoing, or if the data should now be classified as 'post event'. Thus, rainfall nowcasts are used to forecast the future event duration and intensity and so determine a category.

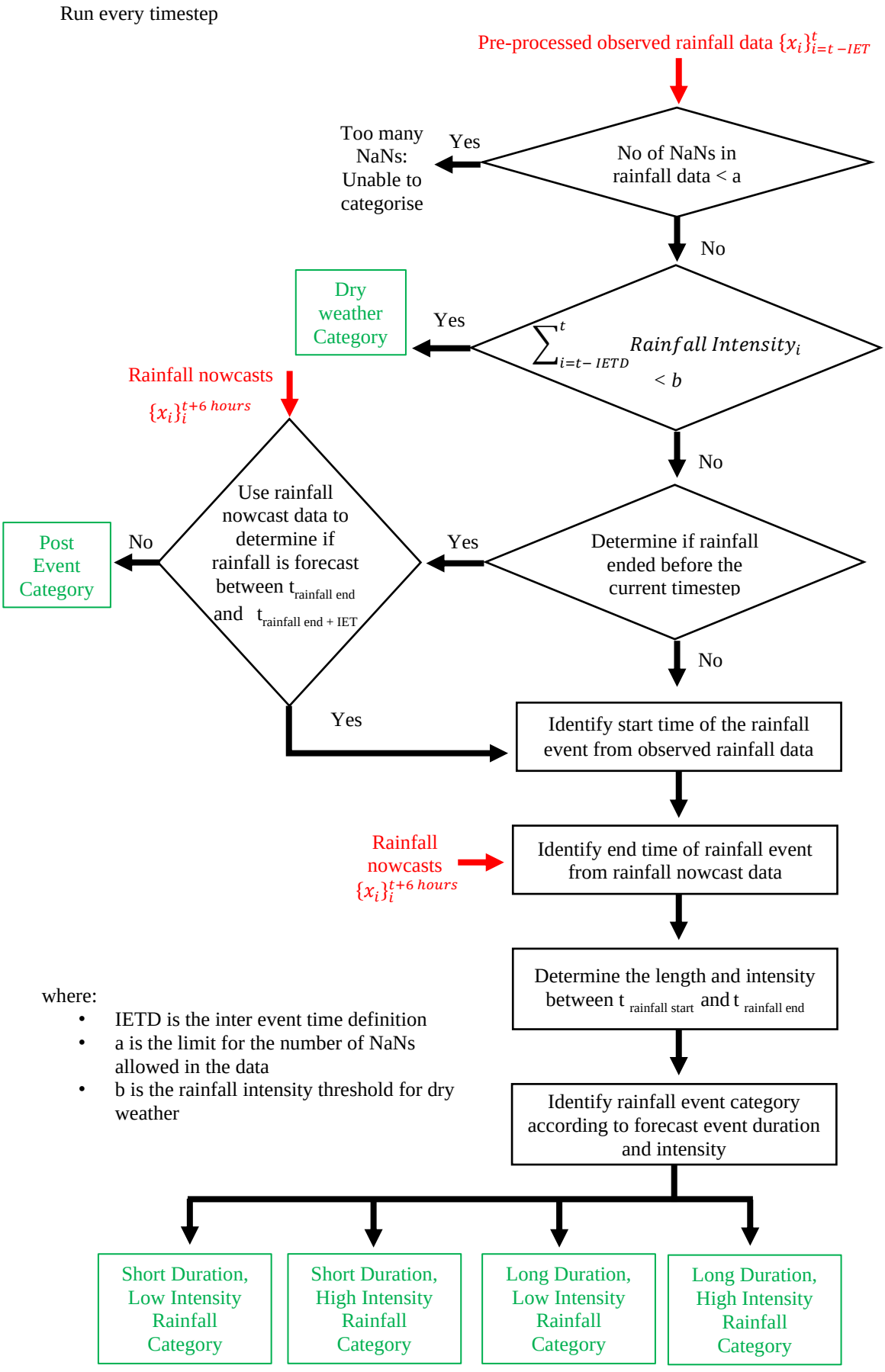
A flow chart describing the real time categorisation methodology is presented in Figure 4-8. Starting with the pre-processed rainfall data, a window of data from t_{IET} to t_0 is analysed. A heuristics-based procedure is first applied to discard the data if too many missing values (i.e. NaNs - Not a Number) are present. The value of a, the limit for the maximum number of NaNs allowed, has been set as

$$a_t = \begin{cases} 10\% & if \sum_{i=\mathrm{t-IET}}^{t} R_i > 0 \\ 20\% & otherwise \end{cases}$$

where *R* is rainfall intensity.

The value of *a,* if the cumulative rainfall is above 0, was set as 10% after a relevant sensitivity analysis (not shown here) which demonstrated that the absence of 10% of the data did not significantly affect the categorisation of the data. The value of *a* is increased to 20% if all the non-NaN datapoints are 0. This is because, as already stated, the majority of a rainfall dataset is 0 mm/hr, and non-rainfall values generally occur consecutively together in a 'pulse'. Therefore, the likelihood than the missing points are 0 when the cumulative rainfall intensity is 0 is very high. *a* is defined as a percentage, rather than a number, to account for changes in the data sampling rate or for a change in the value of the IETD.

If the number of missing values in the data is below the threshold, the rainfall category can then be determined using the process as described in Figure 4-8, using forecast rainfall data to predict the future rainfall behaviour. It is possible that the end time of an event cannot be identified, as the event continues beyond the 6 hour nowcast available, however, this is not an issue, the exact length of the event is not required and it is categorised into the long duration category.



where:

- IETD is the inter event time definition
- a is the limit for the number of NaNs allowed in the data
- b is the rainfall intensity threshold for dry weather

Figure 4-8 Flowchart of the process for categorising data in real time.

This categorisation process is run at each timestep in real time to identify the category of the current timestep. Rainfall nowcasts are constantly updated in real-time – at each timestep the latest forecast is received from the Met Office, consisting of 24 forecasts ranging from 15 minutes to 6 hours ahead at 15-minute intervals. It is therefore possible that the forecast duration and intensity of an event based will change as the event is ongoing and the rainfall nowcasts are updated, and this in turn could result in the assigned rainfall category of an event changing (e.g. a rainfall event is identified as category (iv) at time t and category (v) at time t+1). However, in practice the classification of the events over time is observed to be fairly stable, and categories rarely change.

The results of the 9 year rainfall dataset shown in Figure 4-5 of the rainfall events, categorised according to the above thresholds is presented in Figure 4-9(a). Each point on the graph represents an individual rainfall event, plotted according to the length of the event and the average intensity of the event (i.e. the average of the

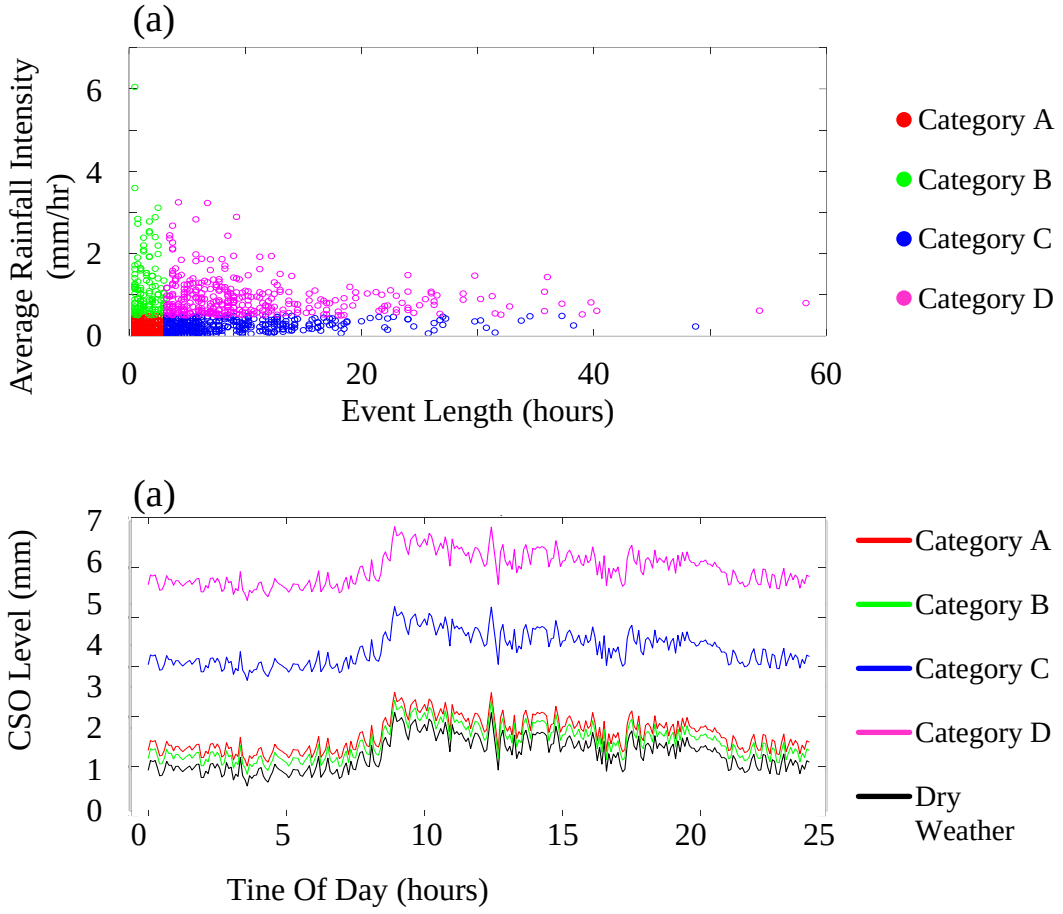


Figure 4-9 (a) Categorised rainfall events, and (b) mean daily level per rainfall category, for a CSO site.

radar rainfall intensity data collected during the duration of the rainfall event). Using this data, the mean CSO level for each timestep of the day could then be calculated for each rainfall category, displayed in Figure 4-9(b). The difference in CSO level between the different categories can clearly be seen, indicating that the various categories are distinct.

4.4 Statistical Trend Based Evidence Generation Module

The next stage of the system is blockage detection. The Statistical trend-based evidence generation module is the first module designed to detect blockage events by identifying unusual trends in the CSO chamber level in real time. The module determines if the incoming level data lies inside a data envelope of 'normal behaviour', defined by statistical analysis of the CSO system.

Historic level and rainfall data during normal operations (i.e. time periods containing no blockage events) are used to calculate the expected levels in the CSO chamber for each of the 6 rainfall categories during the set-up mode of the EDS. During real time running of the system the most recent processed and categorised data are retrieved from the processed time series database for a time window of *d* hours. Control charts are then used to monitor the window of CSO level data to detect any unusual behaviour and, if deviations from the normal operating performance are identified, determine if they are significant enough to provide evidence of a blockage event.

4.4.1 Control Chart

Control charts are one of the most prominent Statistical Process Control (SPC) techniques and are widely used as an efficient means to monitor manufacturing and business processes over time. A typical example of a control chart is shown in Figure 4-10. The chart utilises upper and lower control limits to determine if, statistically, a process is behaving as expected or if it is 'out of control', i.e. the process is unpredictable as a fault has occurred. The limits of the control chart are typically drawn at three standard deviations from the centre line (the statistical mean of the process). Control charts are designed to quickly identify out of control behaviour, allowing necessary corrective action to be taken before a large quantity of nonconforming products are manufactured (or in this case to identify

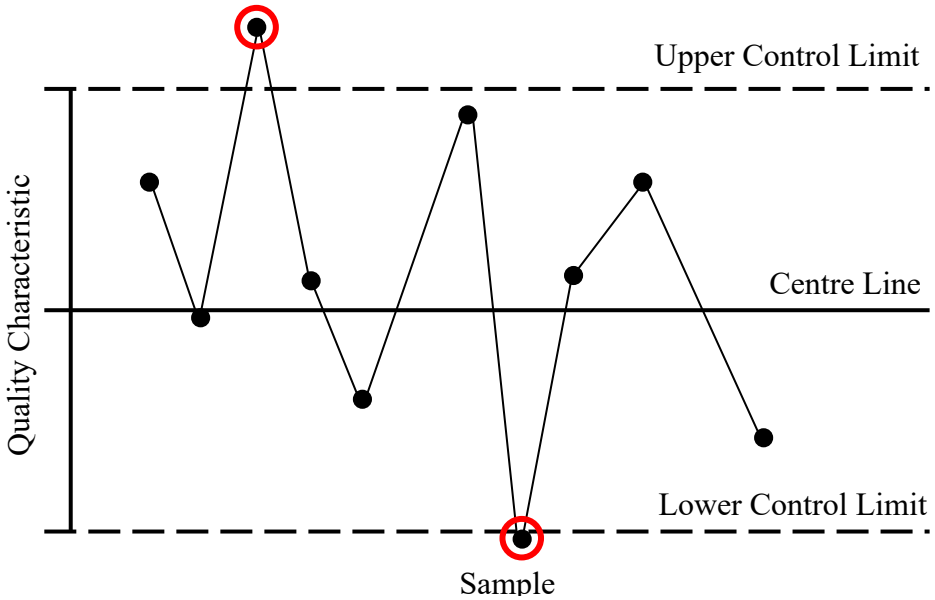


Figure 4-10 Example control chart.

and manage sewer blockages before flooding or other undesirable consequences occur).

Standard control charts are very efficient at detecting large and fast deviations from the process average; however they are generally insensitive to small changes. The statistical sensitivity of the chart can be improved by using supplementary SPC rules that enhance the charts' effectiveness. These SPC rules are intended to recognize a sequence of unnatural patterns, increasing the charts' sensitivity to out-of-control events when compared to the use of the standard single 'three standard deviations' rule. Several different rule sets have been suggested (e.g. Page, 1955; Roberts, 1958; Bissell, 1978; Wheeler, 1983). However, the most popular and well-established are the Western-Electric (WE) rules (Western Electric Company, 1958) , which were developed by the manufacturing division of the Western Electric Company in 1956. These rules have since become a well-accepted standard in the industry. The Western Electric rules are defined as follows:

1. Any single data point falling outside the [$\mu - 3\sigma$, $\mu + 3\sigma$] range
2. Two out of three consecutive data points falling beyond the [$\mu - 2\sigma$, $\mu + 2\sigma$] range
3. Four out of five consecutive data points falling beyond the [$\mu - \sigma$, $\mu + \sigma$] range, on the same side of the centreline
4. Eight consecutive points falling on the same side of the centre-line

where μ is the mean of the observations and σ is the standard deviation.

Data satisfying any of these conditions indicates that the process is out of control and justifies investigation.

The Western Electric control rules were constructed for the statical trend based module using historic CSO level data. As explained above, when operating in real time the incoming CSO level and rainfall data are categorised into 6 categories, according to the current rainfall intensity and duration. This is performed to better customise the blockage detection methodology to the current rainfall characteristics. The control chart limits were thus calculated separately for each of the 6 rainfall categories. This was achieved by computing the mean and standard deviation values of the CSO level data separately for each rainfall category and then calculating the respective control limits in the usual way.

When compared to the single three sigma run rule, the use of WE control rules were demonstrated to significantly increase the probability of detecting blockage events and decrease the detection time of the statistical analysis-based module. However, the false positive rate was unacceptably high. Thus, a modified version of the WE run rules was designed in order to maintain this high level of blockage detection whilst attempting to minimise the number of false alarms produced.

4.4.2 Modified Western Electric Rules

The modified western electric run rules are defined as follows. Various rules are presented in Table 4-1. Rules 1 to 4 consist of the original Western Electric run rules. Rules 5 and 6 were added to accommodate gradually forming blockages, which cause CSO levels to increase slowly over a long period of time. Analysis indicated that these events were frequently overlooked by the original rules, as they require a large deviation from normal levels over a small number of timesteps.

For each run rule *i* the control chart limit *L* for time step *t* is then defined as follows:

If Dry Weather (i.e. rainfall category (1))	$L_{i,dry,w,t} = \mu_{dry,w,t} + M_i * N_{dry} * \sigma_{dry,w,t}$
If Wet Weather	$L_{i,c} = \mu_c + M_i * N_c * \sigma_c$

(i.e. rainfall categories (2) to (6))

where μ_c and σ_c are the mean and standard deviation of the historic CSO level for the current rainfall category c, $\mu_{dry,w,t}$ and $\sigma_{dry,w,t}$ are the mean and standard deviation of the historic CSO level for the dry weather rainfall category at timestep t, w denotes if the timestep is if the timestep is a weekday or weekend. M_l is the constant multiplier defined for each run rule (e.g. 3 in the standard 3-sigma control chart) and N_c is an additional multiplier determined for each rainfall category c.

An extensive sensitivity analysis, described in the following section, was conducted to determine the best combination of run rules and corresponding parameter N for each rainfall category. During real-time running of the system the rainfall category of the current timestep is identified and the appropriate control limits calculated. If the measured level falls outside the limits and any of the rules are satisfied, the evidence of a blockage is inferred.

Western electric rules generally apply upper and lower thresholds. However, as explained above, the system has only been developed for downstream blockages

Table 4-1 Modified Western Electric rules for Shewhart control charts.

SPC Run Rule	Rule Description	Constant Multiplier (M)
Rule 1	1 out of 1 consecutive discrepancies fall outside the defined control limits	5
Rule 2	2 out of 3 consecutive discrepancies fall outside the defined control limits	4
Rule 3	4 out of 5 consecutive discrepancies fall outside the defined control limits	3
Rule 4	8 out of 8 consecutive discrepancies fall outside the defined control limits	2
Rule 5	15 out of 15 consecutive discrepancies fall outside the defined control limits	1
Rule 6	25 out of 25 consecutive discrepancies fall outside the defined control limits	0.5

(which cause an increase in CSO level), due to the rarity of identified upstream blockage events. Thus, only upper limits have been used here.

The SPC rules are defined separately for wet and dry weather timesteps as during dry weather the CSO level is highly dependent on the time of day and the day of the week (the level in a CSO during dry weather exhibits a strong diurnal pattern). This is accounted for in the control chart limits during dry weather by calculating μ and σ for each time step of the day and for weekends and weekdays. In this way the limits are able to adapt to the time varying behaviour of the system. During wet weather the level is dependant primarily on rainfall characteristics and the diurnal pattern is not significant. It is therefore not necessary to consider the timestep or the day of the week when calculating the control chart limits.

The values of μ, σ and N are determined separately for each rainfall category so that the control limits are tailored to the rainfall at the current timestep, therefore increasing the likelihood that blockage events are detected quickly and reducing false alarms. For each run rule a range of different values of N were tested. The combination of rules and N_c parameters that raised a low number of false alerts whilst still maintaining a fast detection time were selected. The selected parameters are designed to be generic, i.e. they can be applied to any CSO without further analysis as they are based only on the μ and σ of the CSO level data.

4.4.3 Selection of Modified Rules

The modified rules were selected following an extensive sensitivity analysis to determine the optimal combination of run rules and the corresponding parameter N for each rainfall category. These analyses included an ROC curve, the partial area under the curve (PAUC) (Ma et al., 2013), and two methods to determine the optimal cut-off point of the ROC curve; the Youden index (Youden, 1950), and a cost-based approach (Zweig & Campbell, 1993). The analyses were performed using CSO data from real blockage events from 15 different CSO chambers, containing 15 blockage events (both gradual and suddenly forming) and located in both urban and rural catchments. A total of 9 years of CSO level and rainfall data was analysed. The selected parameters were then tested on a

further 5 validation (i.e. unseen) CSO sites. The detailed results of these tests are presented in Appendix A.

These analyses were designed to identify run rules which produced a low false positive rate whilst also ensuring a high true positive rate. In particular, the cost-based approach weighs the benefits of true positives against the harm of false positives. In addition, the blockage detection time for the different rules was analysed, to ensure that decreasing the false alarm rate did not significantly increase the detection time of the system.

It should be noted that these parameters, and indeed all the parameters selected for the EDS, were chosen based on careful analysis of the data, and are designed to produce a good performance for all CSO types. However, when the system is deployed by a utility they will prioritise certain features e.g. require a very fast detection time or an extremely low number of false positives. Ultimately. The optimal thresholds will be based on the circumstances in which the system is being employed and on the operator's desires and expertise. In these cases, the parameters can be easily changed, and the analyses presented in Appendix A can be used to facilitate making an informed selection.

The selected run rules and corresponding parameters for each rainfall category are displayed in Table 4-2. The heavier rainfall classes (C and D) were found to

Table 4-2 Selected modified Western Electric rules for the statistical analysis-based detection module.

Category	Selected SPC rules	N
Dry Weather	2-4	3
Rainfall A	2-4	4
Rainfall B	2-4	4
Rainfall C	2-4	5
Rainfall D	3-5	6
Post Event	2-5	4

Category	μ	σ
Dry Weather	See graph	See graph
Rainfall A	151	17
Rainfall B	182	49
Rainfall C	195	101
Rainfall D	338	234
Post Event	135	67

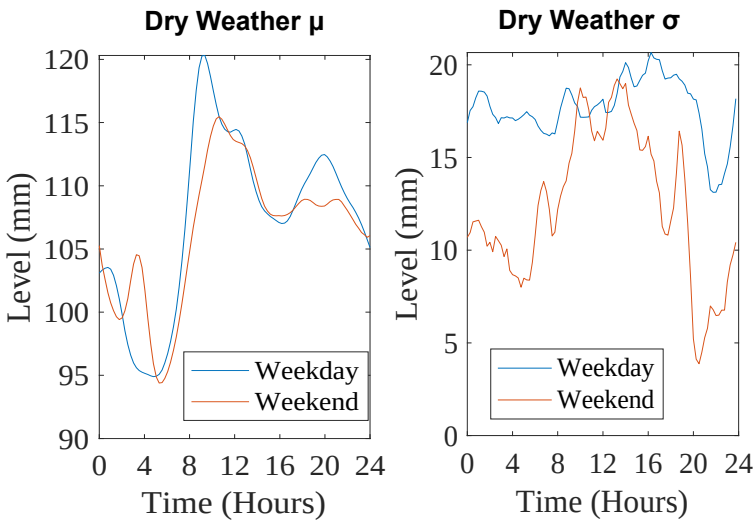


Figure 4-11 Example of the μ and σ values used to generate control limits.

produce a higher number of false events, despite higher control rule boundaries generated due to the higher values of *μ* and *σ* calculated for these classes. Thus, higher values of N were selected for these classes and rules were selected which required waiting for additional timesteps before raising an alarm.

Figure 4-11 shows an example of the statistical values used to generate the control limits for a CSO site for the 6 rainfall categories. The boundaries were calculated using 6 months of CSO level and rainfall data (from July 2016 to January 2017). The difference in μ and σ for the different rainfall categories are evident. The control limits generated using these values will thus be distinct for each rainfall category.

4.5 EANN Discrepancy Based Evidence Generation Module

The EANN discrepancy-based module is the second module designed to identify blockage events. This module aims to detect blockages by analysing the CSO level forecasts produced by the EANN model described in Chapter 3. The EANN model has been trained to forecast normal levels in the CSO chamber, assuming no blockage event. When a blockage occurs the level in the sewer system exhibits abnormal behaviour and so the EANN model is unable to produce accurate forecasts. As a result, the discrepancy between the model prediction and the measured level increases. By continually monitoring this discrepancy in

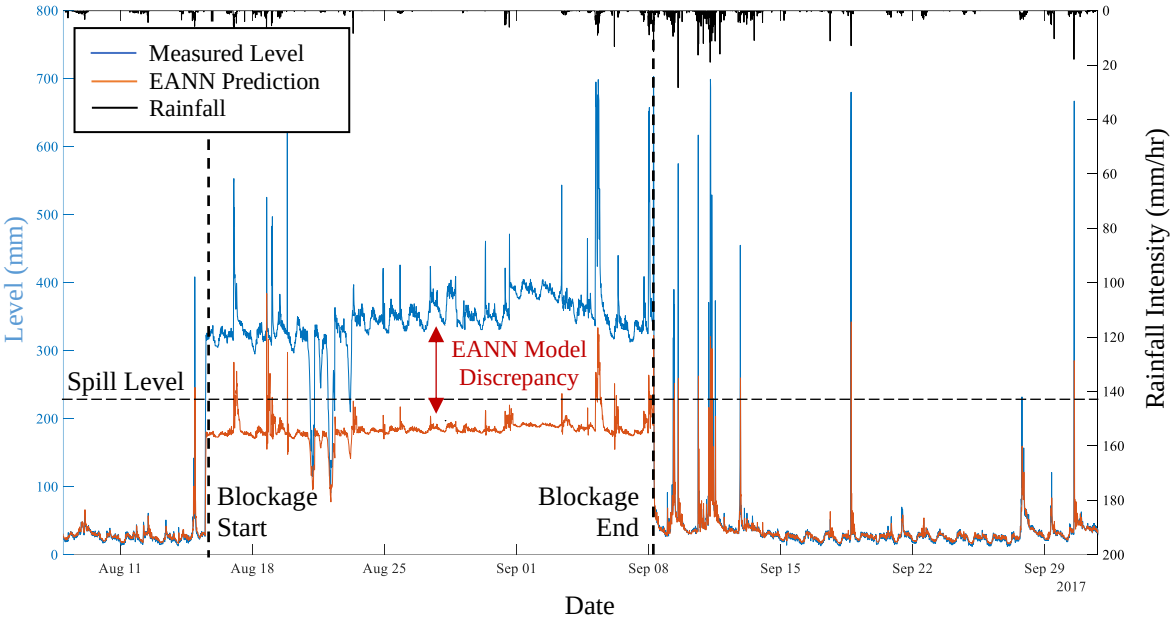


Figure 4-12 Example of the EANN model prediction vs measured CSO level during a sudden blockage event.

real time deviations from normal behaviour can be identified and analysed to infer the presence of a blockage event.

As described in Chapter 3, an EANN model has a greater prediction accuracy than an ANN model with parameters selected via trial and error. In addition, as an evolutionary strategy algorithm is employed to automatically select the optimal artificial neural network structure and parameter set, no significant human input is required to construct the model and the network is tailored specifically for the CSO locations in question. Therefore, the application of an EANN model here compared to a simple ANN is advantageous.

An example of the discrepancy between the EANN model forecast and the measured CSO level during a sudden blockage event is displayed in Figure 4-12. As can be seen from this figure, the blockage occurs in August, causing a significant increase in CSO water level, lasting one month. Although the EANN model was able to forecast the change in level it failed to capture the magnitude of the increase and there is a large discrepancy between the measured and forecast results.

The model discrepancy during gradually forming blockages is generally more difficult to detect as the change in level over time is smaller and the EANN model has more time to adapt to the change. An example of a gradual blockage is shown

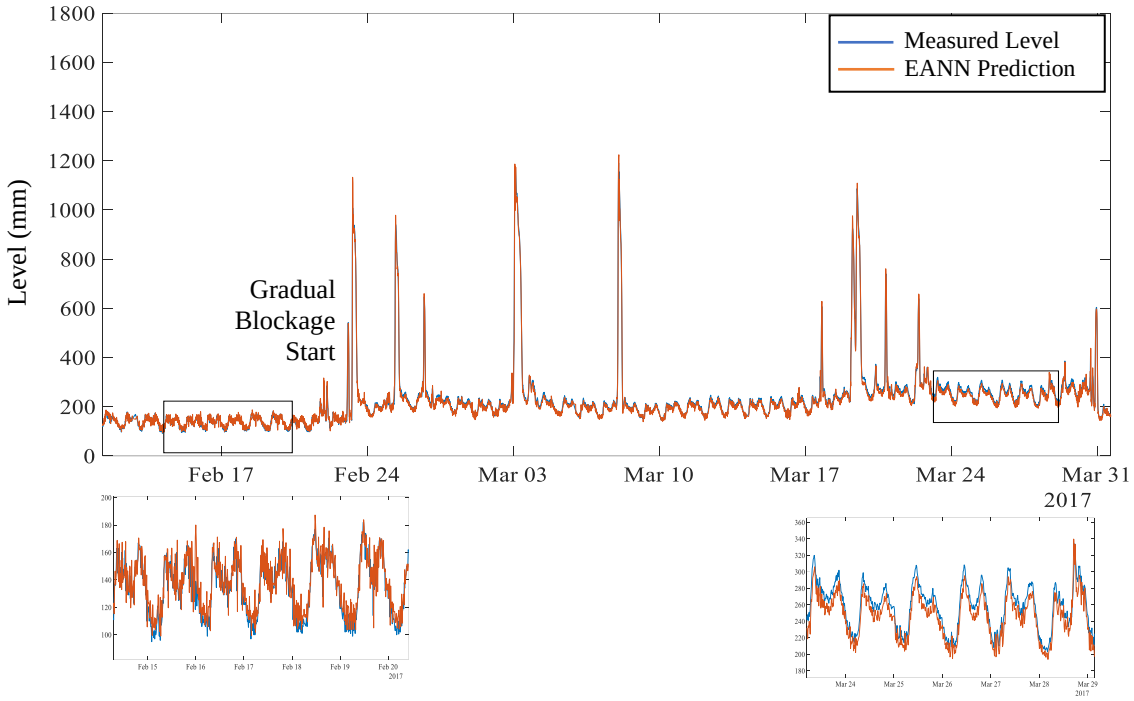


Figure 4-13 Example of the EANN model prediction vs measured CSO level during a gradual blockage event.

in Figure 4-13. The blockage begins on the 22nd February approximately and lasts for one month, before it is removed on 31st March. As can be seen, the error between the forecasts and measured level increases due to the blockage event, however the discrepancy is less obvious.

As with the statistical trend-based analysis module, the EANN model discrepancy is analysed to determine if it exceeds pre-defined limits using modified Western Electric control charts. The charts continually monitor and asses the EANN model discrepancy in real time. The discrepancy is defined here as $x_{EANN,t} - x_{obs,t}$, where $x_{obs,t}$ is the observed sensor level and $x_{EANN,t}$ is the EANN model prediction at time t.

The control chart limits $L_{i,c}$ for run rule i and rainfall category c are then defined as:

$$L_{i,c} = \mu_{EANN,c} + M_i * N_c * \sigma_{EANN,c}$$

where $\mu_{EANN,c}$ and $\sigma_{EANN,c}$ are the mean and standard deviation of the historic EANN discrepancy for rainfall category c and M_i is a constant multiplier defined for each run rule i.

Unlike the statistical analysis module, the limits do not need to be calculate separately for dry and wet weather, as during dry weather the model discrepancy is not affected by the day of the week or the time of day. The accuracy of the EANN model predictions are significantly influenced by rainfall intensity, however. The EANN model performs extremely well during dry weather, thus the model discrepancy is extremely small. During heavy rainfall the CSO level is much more changeable and sudden in nature, and the model discrepancy increases. Therefore, as with the statistical analysis module, calculating the μ and σ separately for each rainfall category and tailoring the control chart limits has the potential to improve the system.

The modified Western-Electric rules and the corresponding parameters were identified using the sensitivity analysis tests utilised in the selection of the statistical analysis module run rules (see Section 4.4.3) and were performed on the same set of blockage events. Further detail on these tests is presented in Appendix A. The selected rules for the EANN discrepancy based module for each rainfall category are displayed in Table 4-3

Table 4-3 Selected modified Western Electric Rules for the EANN discrepancy-based blockage detection module.

Category	Selected SPC rules	N
Dry Weather	3-6	1
Rainfall A	3-6	0.25
Rainfall B	3-6	0.5
Rainfall C	4-6	2
Rainfall D	4-6	3
Post Event	3-6	2

4.6 Rainfall Categorisation Evaluation

In order to assess the performance using the rainfall categorisation method described in Section 4.3 (i.e. categorising the data into 6 rainfall categories), the performance of the statistical analysis and the EANN discrepancy based analysis

evidence generation modules were compared when employing two other rainfall categorisation methodologies:

i. 1 Rainfall Category Methodology: This methodology uses one set of control rules for all weather types. This approach is very simple and has the advantage that data cannot be miscategorised due to inaccurate forecast rainfall data. However, the rules must accommodate all weather types, from dry weather to heavy rainfall, and therefore will necessarily be very broad.
ii. 3 Rainfall Category Methodology: The 3 rainfall categories used here consist of Dry weather, Wet Weather and Post Event. This methodology was included to determine if categorising the rainfall into 6 classes results in insufficient data in certain classes.

It was decided not to test the performance using more than 6 rainfall categories as dividing the dataset into more classes generally will not guarantee enough rainfall events per class to create adequate sewer level profiles. This was demonstrated using an analysis of rainfall and level data sets from 30 CSO sites to investigate how increasing the number of rainfall categories affects the number f rainfall events identified in the data. The results are shown here for one site only due to space limitations. Figure 4-14 compares the results of the data classification when using (i) 6 categories (i.e. the selected methodology), (ii) 8 categories consisting of 2 duration thresholds and 1 intensity threshold, (iii) 8 categories consisting of 1 duration threshold and 2 intensity thresholds, and (iv) 11 categories. The post event and dry weather categories are the same for the 4 different classification methods, and so the results for these categories are not shown in the figure. Six months of data was analysed, from January 2016 to June 2016. The figure presents the total number of rainfall events identified in the data and allocated to each category, and the mean CSO level produced during these rainfall events.

When divided into 6 categories (Figure 4-14(i)) it can be seen that there is a satisfactory number of events in each class – the lowest number of events in any one class is 26. In addition, the mean CSO levels for each category are distinct from each other – indicating that there is meaningful difference in the level response caused by the rainfall events in the various classes. Compare this to

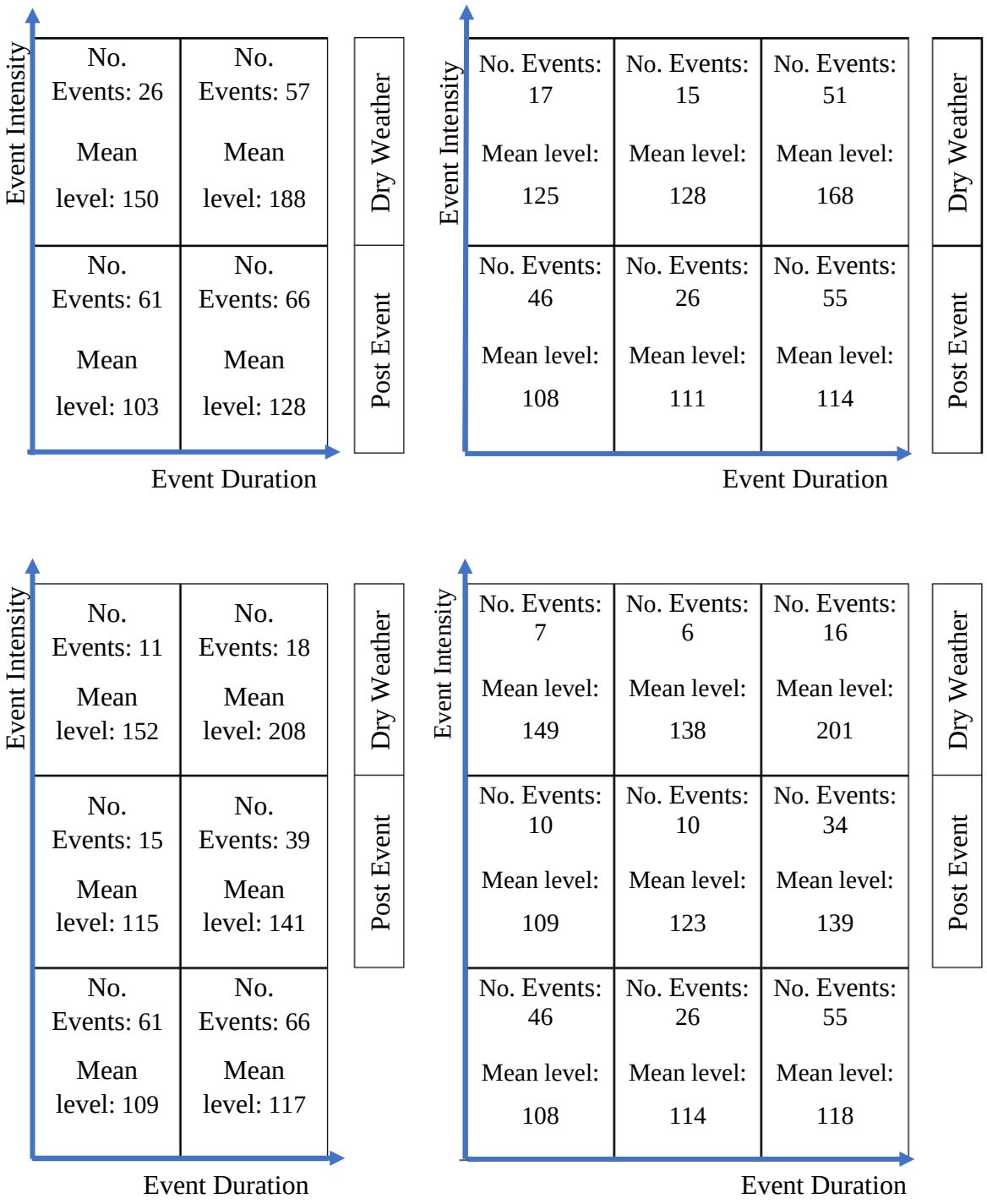


Figure 4-14 Comparison of the number of rainfall events and the mean CSO level (mm) when dividing rainfall into 4, 6 and 9 wet weather categories. The dry weather and post event categories are the same for all schemes.

the results produced when dividing the data into 8 categories (Figure 4-14(ii) and (iii)); the number of rainfall events in several categories are low (often below 20), and the mean CSO level for several classes is similar, indicating that certain

classes are redundant as they contain events which produce similar level responses.

This is even more evident when dividing the data into 11 categories - one class only has 6 events and many of the mean CSO levels are within a few millimetres. This example presents 6 months of data and obviously the lack of rainfall events for certain classes is exacerbated when using a shorter dataset. This problem was identified in all 30 analysed datasets, indeed in some cases the least populated categories only had 2 identified events. Thus, it was decided not to consider methodologies using more than 6 categories, in addition to the 1 rainfall category and 3 rainfall category methodologies described above.

The WE run rules and parameters for the 1-category and 3-category methodologies for the statistical analysis-based module and the EANN discrepancy-based module were selected using the same sensitivity analyses performed for the 6-category methodology, as described in Appendix A. The modules utilising the 3 different rainfall categorisation methodologies were then tested on 7 unseen validation blockage events from different CSO sites. Four months of level and rainfall data were used to train the system for each CSO site, and a minimum of 3 months of data was used for validation.

In general, the more data used to train an EANN model the better, however it was found that 4 months of data still provides very good forecast results. Unlike a physical model, which can be applied to any site it is calibrated for, a data driven model requires a historical dataset before it can be implemented. Requiring only a few months of CSO level data means that a wastewater utility can apply the system quickly after installing a level monitor in a CSO. The time periods of the training and testing datasets used here vary by site, as the blockage events occurred at different times and the level monitors were installed in the CSO chambers on different dates.

The results of the modules' performance are presented in Table 4-4. The different rainfall categorisation approaches were evaluated according to the number of blockages correctly detected, the rate of false positives and the blockage detection time. Overall, it was determined that the 6-category methodology produced the best results both detection modules. Regarding the statistical analysis module, the 6-category approach detected the highest number of

Table 4-4 Comparison of the Statistical Analysis and EANN discrepancy-based module results when utilising different rainfall categorisations.

		1 Category Methodology			3 Category Methodology			6 Category Methodology		
	Blockage Event	**False Alarm Rate**	**Total No of False Alarms**	**Blockage Detection Time (Hours)**	**False Alarm Rate**	**Total No of False Alarms**	**Blockage Detection Time (Hours)**	**False Alarm Rate**	**Total No of False Alarms**	**Blockage Detection Time (Hours)**
Statistical Analysis Based Module	**1**	2.0×10^{-4}	2	Not Detected	0	0	1.75	0	0	1.75
	2	0	0	23.35	0	0	190.75	0	0	18.75
	3	0	0	25.25	0	0	25.25	0	0	3.00
	4	0	0	6.25	0	0	Not Detected	0	0	3.00
	5	0	0	23.75	5.70×10^{-4}	4	65.75	1.43×10^{-4}	1	12.50
	6	1.3×10^{-4}	1	23.00	5.34×10^{-4}	4	4.25	4.01×10^{-4}	3	4.75
	7	0	0	Not Detected	0	0	Not Detected	0	0	Not Detected
EANN Discrepancy Based Module	**1**	0	0	Not Detected	0	0	24.25	0	0	39
	2	0	0	8.50	0	0	248.5	0	0	0.75
	3	0	0	22.75	0	0	70	0	0	27
	4	2×10^{-3}	11	3.50	2.7×10^{-3}	25	75.25	3.1×10^{-3}	29	3.5
	5	0	0	Not Detected	1.34×10^{-4}	1	0.75	0	0	4.75
	6	0	0	Not Detected	0	0	Not Detected	0	0	Not Detected
	7	0	0	3.75	0	0	24.25	0	0	39

blockage events (6 out of 7, compared to 5 out of 7 for the other two methodologies) and, for 5 out of 6 events, also attained the shortest detection time. Indeed, all the events were detected in under a day. The number of false positives generated were similar for all three approaches – the 6-category methodology produced 4 false alarms across all the case study sites, the 3-category methodology produced 8 and the 1-category methodology produced 3. Regarding the EANN based module, again the 6-rainfall category approach produced better results overall, detecting 6 out of 7 events with generally low detection times. The 3-rainfall category approach also detect 6 out of 7 events, however the detection times were generally higher. The 1 rainfall category approach produced a low number of false alerts and had a generally low detection time, however only detected 4 out of the 7 events.

4.7 Inference Module

4.7.1 Overview

The inference module constitutes the final processing stage of the event detection system and enables the raising of alarms when a blockage is detected. When applied in real time the EANN discrepancy-based analysis module and the statistical trend-based module are run in parallel, and the results are fed to the inference module. At each timestep the inference module analyses the incoming information from the two systems and determines if there is enough evidence of a blockage event to justify raising an alarm. In the case of an alarm the system also provides additional information that may aid in diagnosis and management of the event.

The inference module is composed of a rule-based inference engine which applies logical rules to the incoming data in order to deduce new information. The engine is forward-chaining (also known as data driven), meaning it starts with the known available data and uses inference rules to extract more information until a goal is reached. The engine executes the inference rules successively in a forward direction until an antecedent (i.e. If clause) is known to be true. The engine then infers the consequent (i.e. the Then clause) (Sharma et al., 2012). This is in contrast to backward chaining, or goal driven inference, which starts

with goals and works backward to determine which facts must be asserted so that the goals can be achieved.

A simple rule-based inference engine has been developed and applied here. More sophisticated inference methods are available, for example neural network based engines, decision trees (Nicolau et al., 2017) and random forests (Prawira Putra et al., 2019). However, these generally require large amounts of event training data, which is not available here. Additionally, the advantage of the rule-based system is that the rules are well defined and easily understandable, the engine can therefore be modified and extended by future users of the system who may not be experts in the field.

4.7.2 Methodology

Figure 4-15 presents a flowchart describing the processing of the inference module. The module receives the evidence from the EANN discrepancy-based analysis are applied. If either module presents evidence at the current timestep that a blockage has occurred past data is then analysed to determine if there is enough evidence to raise an alarm. A window of n past timesteps are retrieved from each module and analysed to determine if the total number of timesteps containing evidence of a blockage event is over pre-specified threshold m. If all these rules are satisfied a blockage alert is generated. Additional rules were considered, such as requiring the abnormality of the level behaviour to increase over the considered timesteps or requiring both detection modules to generate evidence of unusual behaviour before raising an alarm. However, these were not demonstrated to improve the system's results and so were not utilised.

Initially it was assumed that the EANN analysis-based module would be superior at detecting sudden blockages and the statistical trend-based analysis module would be superior at detecting gradually forming blockages. The evidence from each module could therefore be weighted by the inference engine according to the blockage event type detected. However, the results did not demonstrate that either module was more affective at detecting sudden or gradual blockages. Therefore, the evidence from both modules is weighted equally in the inference engine.

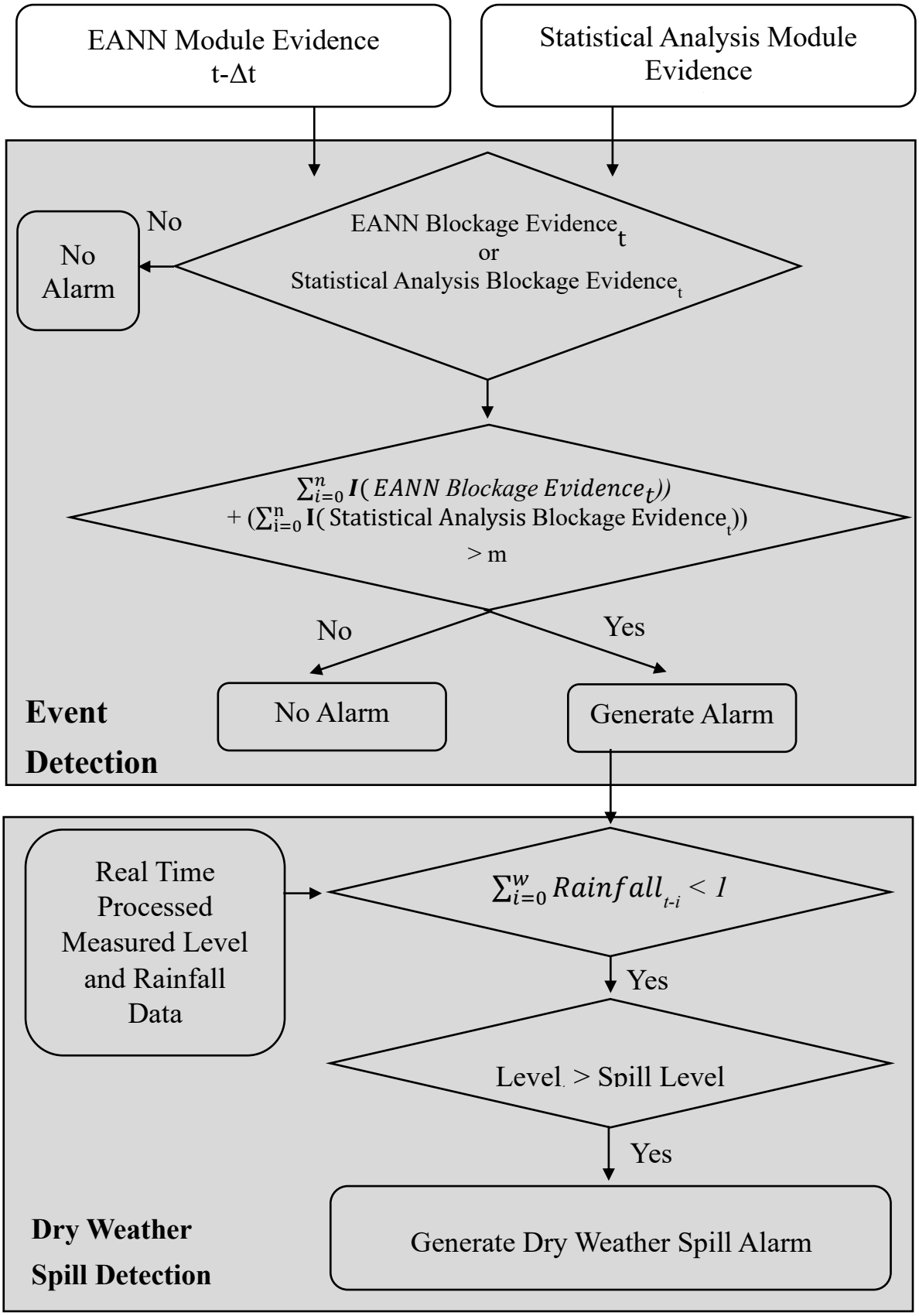


*Figure 4-15 Inference module methodology. **I** is an indicator function that takes value 1 when blockage evidence happens and value 0 when it does not, and n, m and w are parameters.*

The inference module applies a temporary alarm suppression period after every initial alarm is raised which suppresses all subsequent alarms for a specified period of time. This avoids raising unnecessary alarms for the same blockage event. The suppression period was implemented to prevent alarm overloading from multiple alarms. Overloading can make it difficult for the operator to adequately assess all the alarms, and increased stress can lead to poor judgments. The addition of the suppression period also has the benefit of significantly reducing the false positive rate.

It was hypothesised that the suppression period could also result in the system failing to detect genuine blockage events or in an increase in detection time if a blockage occurred during this suppression period. However, an analysis of the effect of deploying the suppression period, presented in Chapter 6, demonstrated that the period does not have a significant effect on either the true detection rate or the detection time of the system.

In addition to the blockage detection alarm, the inference module also generates an additional alert if a dry weather spill occurs. CSOs are designed to spill during heavy rainfall however unusual events such as blockages can cause CSOs to overflow during dry weather at low flow rates which are out of consent. Dry weather spills are undiluted by rainfall and contain untreated, heavily polluted foul water and as a result dry weather spills can have a significantly more detrimental impact on the surrounding environment.

It was decided to analyse the data for dry weather spill alerts only when a blockage event has been detected. This is because an analysis of the data indicated that dry weather spills which occurred when a blockage event had not been detected were all due to anomalous level data (e.g. logger malfunctions, benching which was not removed sufficiently etc).

Once an alarm has been generated by the EDS and stored in the alarm database the following information is recorded: (i) The event start time, (ii) The CSO level, and (iii) whether an overflow has occurred. The inference module is designed to operate at each timestep. However, if desired it can be modified to run over a user specified time period, e.g. once per day, and then generate a list of any blockage alarms raised in the analysed period.

4.7.3 Selection of Inference Module Parameters

The selection of parameters for the inference engine was performed in a similar manner to the selection of parameters for the EANN analysis and statistical analysis modules, i.e. by analysing the true positive rate, false positive rate and event detection time and selecting the optimal combination of values. Again, the selection of parameters involved a trade-off between true and false alarms.

Different engines were constructed utilising values of *n* and *m* from 1 to 20 and the different combinations were analysed using a grid search method to identify the best set across all the CSO sites. The selected engine utilises the parameters *n = 6* and *n = 5*.

Figure 4-16 compares the false detection rate vs the mean blockage detection time for the various inference engine parameter combinations considered, with the selected engine highlighted. Increasing the evidence required to generate an alarm decreases the number of false alarms although it also increases the average blockage detection time somewhat. In addition, the systems' results are shown when using no inference engine – i.e. generating an alarm immediately if

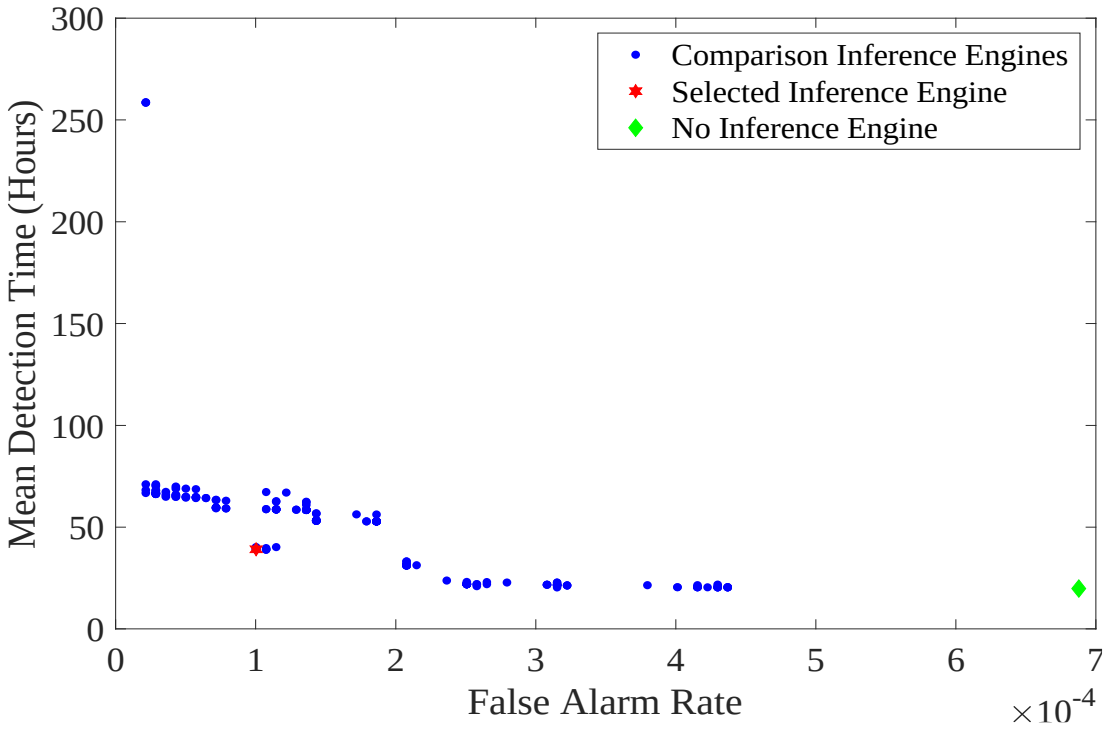


Figure 4-16 False alarm rate vs mean blockage detection time for different inference engine parameters.

either module presents evidence of a blockage event. It is evident that the implementation of the inference engine significantly reduces the false alarm rate and is therefore beneficial to the detection system.

4.7.4 Event Confidence Estimates

An additional functionality was developed to provide the estimated confidence level of a blockage event when an alarm is raised. This information is designed to allow operators to rank alarms and to prioritise any response if a large number of alarms are raised at the same time. The alarms are classified here as low, medium and high confidence. The confidence of the alarm is related to the increase in level and can therefore also be viewed as a measure of the severity of the detected blockage event – i.e. high confidence alarms indicate a greater change in CSO level and are thus is more likely to result in an overflow.

The alarm confidence methodology is designed to estimate the alarm confidence at each timestep a blockage alarm is generated, based on the evidence from the two detection modules. The alarm confidence is determined during the implementation of the inference engine. The inference engine detects blockage events by determining if the total number of timesteps containing evidence of a blockage event from a window of n timesteps of past data is greater than threshold m. The alarm confidence methodology builds on this system, by utilising the number of timesteps containing blockage evidence from this window (called I here) to determine the alarm confidence. Provided a blockage has been identified, a simple approach is implemented whereby a value of $I < t_1$ indicates a low confidence alarm, $t_1 < I < t_2$ indicates a medium confidence alarm, and $I > t_2$ indicates a high confidence alarm. Different pairs of threshold values were analysed designating low, medium and high alarms to identify the best values.

Figure 4-17 presents the results when utilising three different threshold pairs with threshold pair 1 requiring the least evidence to classify an alarm as high confidence and threshold pair three requiring the most. Additionally, threshold pair 3 requires that alarms are only classified as high confidence if the CSO level at the current timestep is greater than 60% of the chamber spill height. The

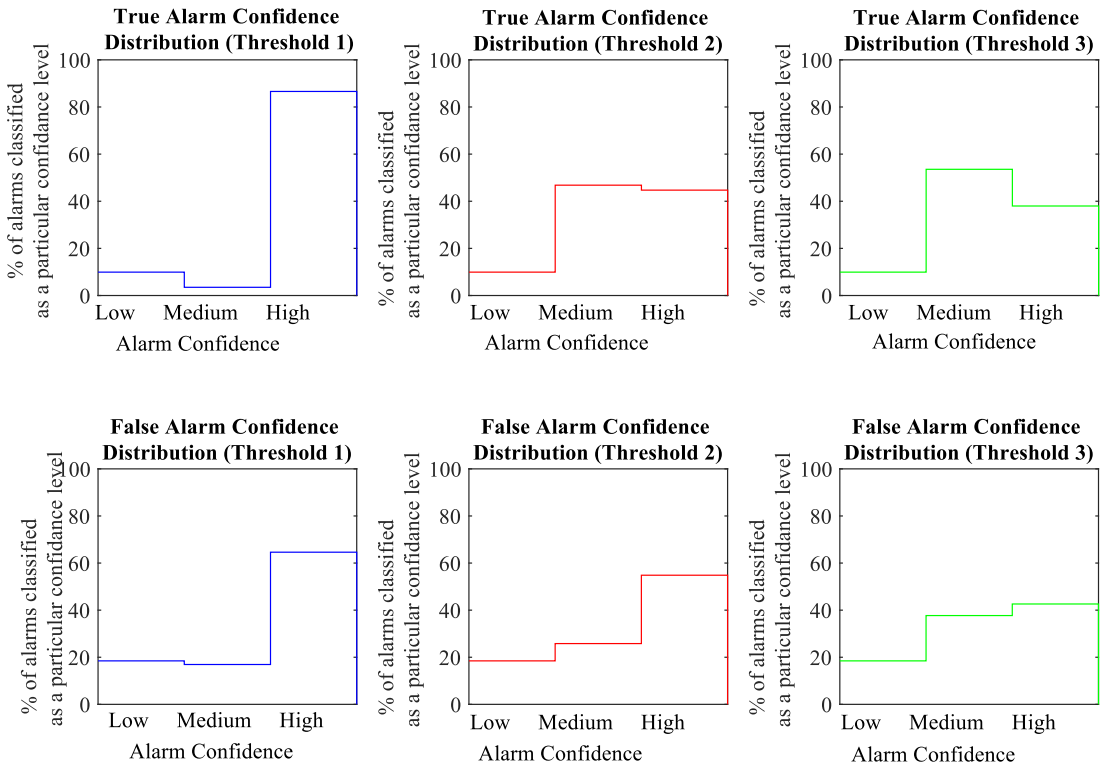


Figure 4-17 Distribution of alarm confidences for three threshold values.

graphs show the percentage of true and false alarms classified as low, medium and high, when using the various threshold pairs, i.e. when applying threshold pair 1, 64% of all false alarms were classified as high confidence.

Based on these results it was decided not to include the confidence estimate functionality in the current detection system. The figures show that when applying threshold pair 1 the majority of true blockage events were correctly classified as high confidence. However, a smaller but still significant majority of false alarms were also classed as high confidence. It was felt that classing false alarms as high probability would erode user confidence in the system and therefore would be overall detrimental. However, when increasing the thresholds required to raise a high confidence alarm (i.e. threshold pairs 2 and 3) there is little difference between the true and false alarm classification distributions of low, medium and high confidence alarms – therefore little extra information is gained.

Therefore, although estimating the alarm confidence is worthwhile, and would be beneficial to a wastewater utility, it requires further work before it can be included in the detection system.

4.8 Summary

This chapter has described a novel wastewater event detection system designed to detect blockages and other unusual events in the vicinity of a CSO chamber in real time. The system also provides additional information designed to aid the user in determining the appropriate response to the blockage event and, in the case of multiple alarms, help in ranking/prioritising the response.

The system consists of a pre-processing and categorisation subsystem, a blockage event generation subsystem, containing 2 event detection modules – an EANN discrepancy analyses module and a statistical analysis module. These modules analyse incoming data to identify evidence that a blockage has occurred. Modified Western Electric rules were designed to improve the effectiveness of these blockage detection modules. The modified rules were demonstrated to improve the false positive rate of the system whilst still generating an acceptable true positive rate detection time. An event inference subsystem is then employed, which combines the evidence from the blockage generation modules and enables the raising of alarms when a blockage event is detected.

During pre-processing of the incoming data, the system categorises rainfall and CSO level data according to rainfall intensity and duration, in order to tailor the system to the current rainfall in the CSO catchment. This was found to improve the blockage detection rate and detection time of the system.

The EDS requires only standard CSO level and rainfall data, datasets which are currently routinely collected by wastewater utilities in real-time, and therefore should not require the installation of any additional monitors in the sewer network. Additionally, the methodology is generic and does not require any physically based hydraulic models. The system can thus be applied to any CSO site without additional work, providing a historic level and rainfall dataset is available for training the system. The system is intended to be automatically recalibrated at predefined intervals, in order to accommodate any changes in sewer network.

The event detection system aims to enable wastewater utilities to automatically detect blockage events in a timely and reliable manner. This will allow utilities to respond promptly to blockage events, therefore reducing potential harm to the

sewer network, and damage to the surrounding environment and properties due to flooding and overflows. The system will also improve the utilities' operational performance and customer service.

Chapter 5: Case Studies for CEANN Level Prediction Model

5.1 Introduction

This chapter presents the results of analyses carried out on case study CSO sites. The aim of these studies is to test, verify and demonstrate the capabilities of the CEANN level prediction model described in Chapter 3 to forecast water levels in a CSO chamber and to provide warnings of upcoming spill events in near real time. The CEANN model is compared with 3 other ANN models, in order to fully evaluate its capabilities. The case studies presented use historical data from level monitors installed in CSO chambers in the United Utilities wastewater network located in the North West of England.

This chapter is organised as follows, first the objectives of the case studies are presented in Section 5.2. Section 5.3 then describes the data used in the case study analyses. Section 5.4 evaluates the performance of the data pre-processing methodology. Next Section 5.5 presents a series of preliminary tests conducted to make design decision for the ANN model. Section 5.6 describes the CEANN CSO level prediction case studies; this includes a description of the CSO sites, the implementation of the CEANN model, and the three comparison ANN models, a description of the results, and an evaluation of models' capabilities. Section 5.7 then evaluates the ability of the CEANN model to predict CSO spill events and to generate alerts. Finally, a summary of the chapter and the main conclusions are given in Section 5.8.

5.2 Objectives of Case Studies

As described above, the aim of the analyses described in this chapter are to test, verify and evaluate the capabilities of the CEANN model for CSO level prediction. More specifically, the objectives of the case studies described here are:

- To investigate the improvements that can be obtained in ANN model performance by using rainfall nowcast data as a model input;

- To investigate the improvements that can be obtained by using an EA to evolve ANN model architecture and input structure, compared to models with architecture and input structure determined manually via trial and error;
- To evaluate the improvements gained by using techniques to overcome data-imbalance;
- To evaluate the ability of the CEANN model to forecast CSO levels in real time and to predict potential overflows at different forecast horizons;
- To evaluate the usefulness of the CEANN model in enabling better management of CSOs by wastewater utilities, in particular the management of upcoming spill events.

5.3 Case Study Sites Data

The studies presented here make use of historical CSO level and rainfall data from various CSO sites. For each site time-series level data (in mm) was collected using an ultrasonic depth monitor, specifically a Cello logger (Technolog, n.d.), located in the CSO chambers. The water level readings were obtained at a uniform temporal resolution of 2 minutes. Observed radar rainfall intensity data (mm/hr) was obtained with a 5-minute temporal resolution and 1 km^2 (i.e. 1x1 km) spatial resolution. Rainfall nowcast data (mm/hr) was obtained with a lead time of 6 hours, a 2x2 km spatial resolution and a 15-minute temporal resolution.

The radar rainfall and nowcast rainfall datasets were obtained through the NIMROD system, an automatic radar analysis and forecast system operated by the UK Met Office. The system receives radar imagery from a network of radar stations in the UK. High-resolution ensemble nowcasts are then generated using the Short Term Ensemble Prediction System (STEPS) which blends extrapolated observations with the most recent high resolution numerical weather prediction (NWP) forecast (Bowler et al., 2006).

All CSO level and rainfall datasets were pre-processed, as described in Chapter 3, to remove anomalies and account for missing data points. During this pre-processing, the datasets were interpolated to a uniform resolution of 15 minutes, to match the resolution of the nowcast data. Regarding the datasets designated for testing of the models in a real-time fashion, it was ensured that the pre-

processing methodology was also applied in a simulated real-time manner, to ensure that the models were tested under real conditions.

Information relating to the case study CSO sites analysed in this chapter are presented in the appropriate sections. All the sites were selected to represent the different catchment and CSO chamber types present in a sewer network. Information is given on the upstream catchment size of the analysed CSOs – this refers here to the sewer network upstream of the sewer chamber and directly contributing to the CSO water level (the upstream catchment identified for a CSO chamber can be seen in Figure 5-6). A small catchment is here defined as a CSO with sewer pipes identified as directly upstream of the CSO chamber encompassing an area less than 1km^2, a medium catchment as one encompassing between 1 and 2 km^2, and a large catchment as greater than 2km^2.

5.4 Data Pre-processing

As stated, all historic and real-time rainfall and level data are first pre-processed according to the data pre-processing methodology described in Chapter 3. This is employed to ensure a continuous, uniform and anomaly-free dataset. The critical component of this methodology is the identification and removal of benching. Benching causes anomalous level readings, which can negatively affect the training of the ANN model. It is therefore important that the pre-processing methodology should effectively remove all erroneous data-points, without incorrectly removing genuine data. The methodology is designed for both historical and real-time datasets.

To test and demonstrate the performance of the data pre-processing methodology it is here applied to data from 3 different CSO sites consisting of: (1) a site containing significant benching, (2) a site containing no benching, and (3) a site containing anomalous data. Information regarding the sites is presented in Table 5-1. All benching was originally identified in the data by a manual visual inspection. The historic data-processing methodology was applied to the first 60% of the datasets, and the real-time methodology applied the remaining 40% the data (i.e. in a simulated real-time fashion).

Table 5-1 Summary table of CSO case study sites' characteristics used to assess data pre-processing methodology.

CSO ID		**1**	**2**	**3**
Location		Lancashire	Wirral	Liverpool
Catchment Type		Urban	Urban	Urban
Upstream Network Size		Large	Small	Medium
CSO chamber	**Height (mm)**	3885	1700	3640
	Spill Level (mm)	200	520	750
Datasets	**Calibration**	31/03/2016 – 23/11/2016	12/3/2016 – 29/9/2016	15/3/2016 – 4/1/2017
	Testing	24/11/2016 - 30/04/2017	30/9/2016 - 15/4/2017	5/1/2017 - 20/7/2017

Significant Benching

The first case study example, CSO site 1, consists of a CSO site located in a residential catchment in Lancashire. The CSO level data exhibits a moderate amount of benching. Figure 5-1 displays the raw level and rainfall data from the site from April 2016 to April 2017. The benching manifests as sharps spikes in level. These are identified as benching, rather than genuine increase in level as (i) they occur at a consistent height, (ii) they do not correspond to rainfall in the surrounding catchment, and (iii) they are only 1 timestep in length. A photograph of the CSO chamber and the level monitor is presented in Figure 5-2; the CSO chamber is confined, and it is likely that the level sensor signal is bouncing off the metal plate rather than the surface of the water.

The results of the data pre-processing are presented in Figure 5-1(b). A total of 137 datapoints were identified as anomalous and removed, 101 from the historic dataset and 36 from the real-time dataset. The results demonstrate that the false readings have been effectively removed, whilst successfully preserving the true water signal.

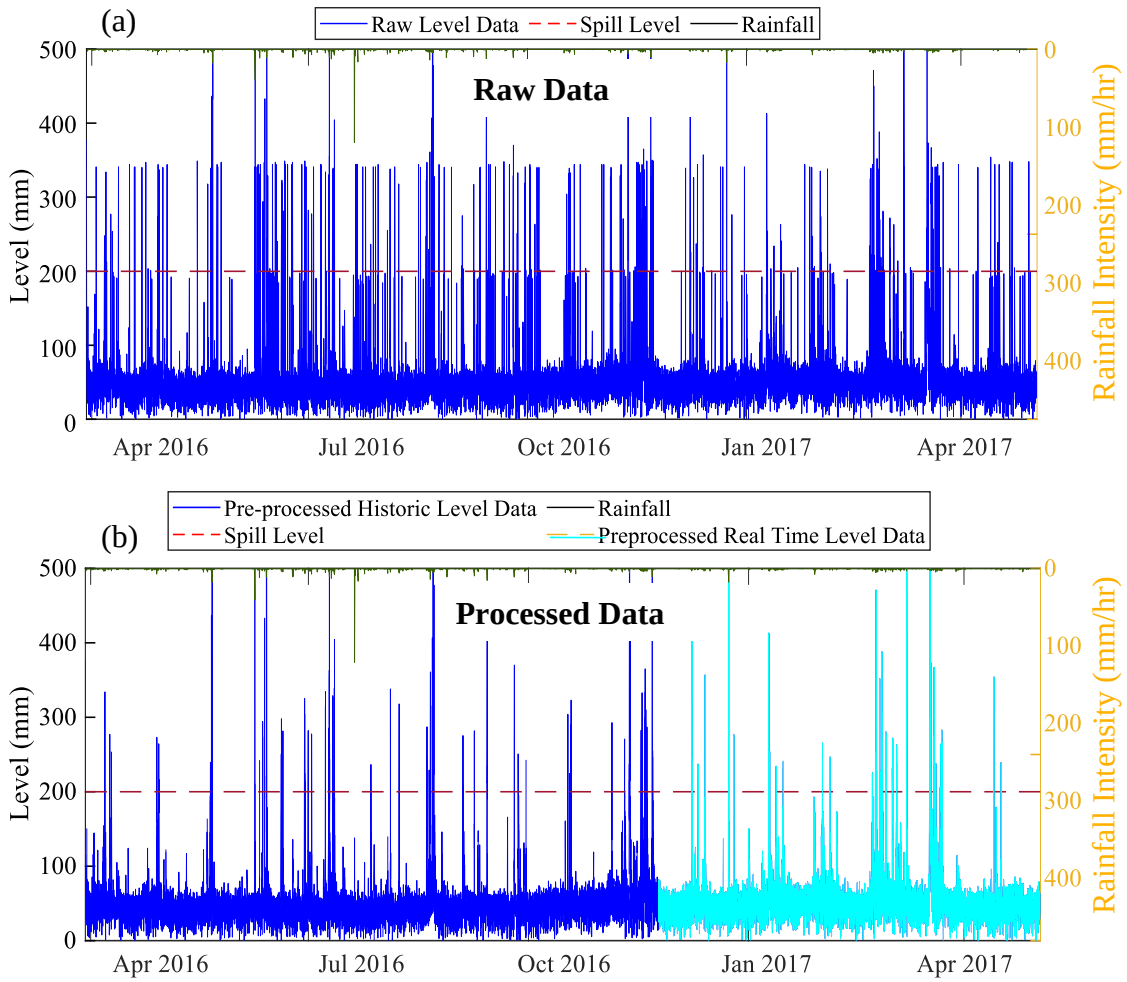


Figure 5-1 (a) Raw level and rainfall data and (b) pre-processed level and rainfall data, from a CSO site exhibiting benching.

Figure 5-2 Photograph of the CSO chamber and level sensor at a site exhibiting benching.

No Benching

The second example evaluates the pre-processing methodology on data from a CSO site experiencing no benching, located in a residential catchment in the Wirral. The level and rainfall data for a 1-year period are displayed in Figure 5-3. The pre-processing methodology correctly determined that no benching was present and therefore no datapoints were removed. A similar effect was obtained for all the other case study sites the methodology was applied to that did not contain benching. Therefore, the methodology can be confidently relied upon not to falsely remove genuine data.

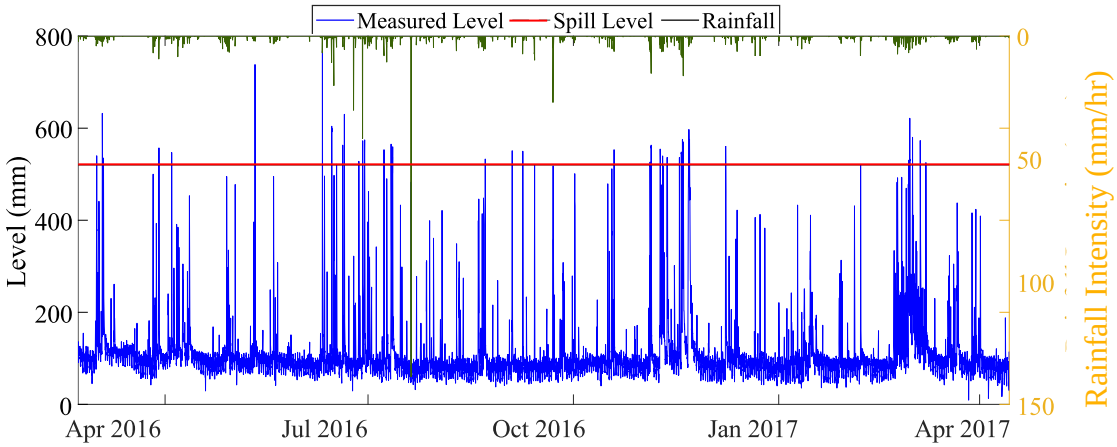


Figure 5-3 CSO level data from a site with no benching.

Anomalous Data

The final example presents data from a CSO site located in a residential catchment in Liverpool. The CSO level data from the site from April 2016 to April 2017 is presented in Figure 5-4. The level data exhibits abnormal behaviour similar to benching, sharp increases in the level occur periodically. However, the characteristics of the readings differ in several key ways. As a result, the pre-processing methodology is unable to identify these anomalies as benching.

It was initially assumed these spikes were genuine increases in level caused by periodic inflows into the chamber, for example from an upstream factory. However, an examination of the CSO chamber revealed that the level sensor in the CSO chamber was positioned close to the inlet pipe, and spray from incoming

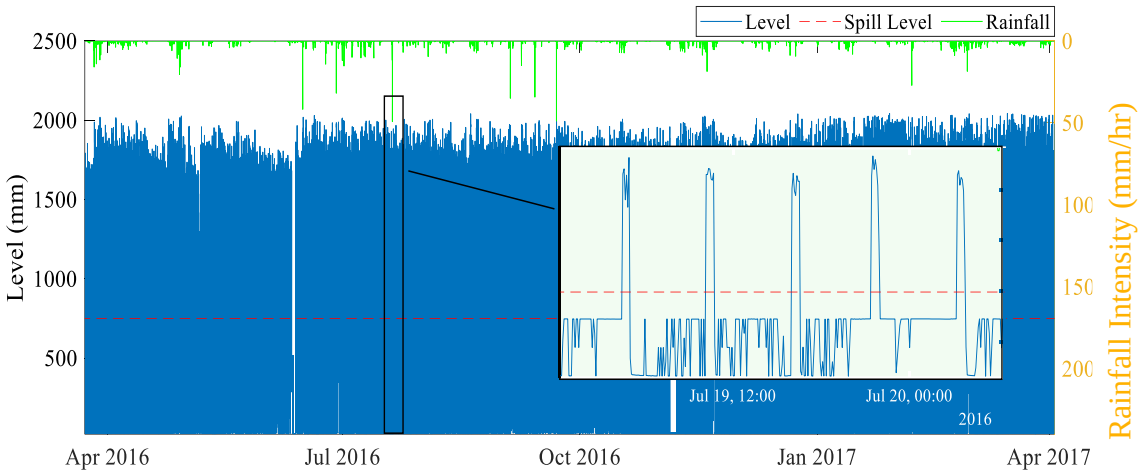


Figure 5-4 CSO level data from a site exhibiting anomalous behaviour.

flow from a pumping station was resulting in false level measurements. A photograph of the CSO chamber is displayed in Figure 5-5. The false readings generated do not have the characteristics of normal benching data - they occur over several timesteps and fluctuate in level, rather than remaining constant. Thus, when the data pre-processing methodology is applied, they are not recognised as benching and so are not removed. This is problematic as these readings are above the spill height of the chamber, and so are recorded as overflows. Regarding the use of the event detection methodology, they could be identified as blockage events resulting in the generation of false alarms.

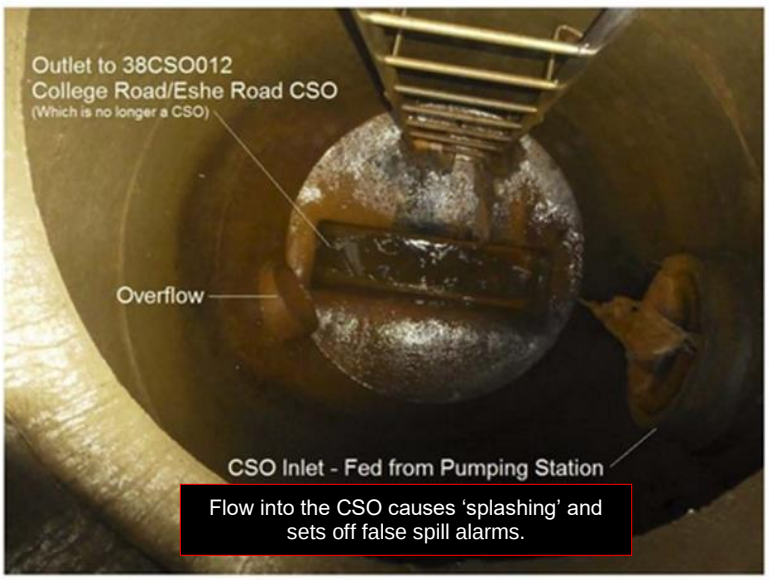


Figure 5-5 Photograph of the CSO chamber at a site exhibiting anomalous data.

The data-pre-processing methodology could be potentially modified in the future to accommodate this type of anomalous data. However, this behaviour is very unusual and so is not further pursued in this thesis.

5.5 ANN Model Preliminary Tests

Before the development of the CEANN model, several preliminary tests were conducted to make various design decisions for the model. These tests were conducted using a conventional ANN model with an input structure and architecture determined via trial and error. The tests were conducted on 10 case study CSO sites, information relating to these sites is presented in Table 5-2. It was ensured that the sites were representative of different types of CSO chamber and catchments.

The results of these preliminary tests are presented here in detail for CSO site 1 only, due to space limitations. This CSO is located in a predominantly urban catchment in the Wirral area of the United Utilities network. The drainage area contains 23,184 properties, serving a population of 51,828 people. The radar rainfall information for this location is provided primarily by a radar station located in Hameldon Hill, Lancashire.

The time series data for the CSO ranges continuously from 27/4/16 to 14/4/17. The data was divided into separate contiguous datasets, for training and testing of the model. The testing dataset consisted of 70%, and 30% was used for testing – an independent assessment of network generalisation on unseen data in a real time fashion. Both datasets contained a representative set of dry and wet weather periods.

5.5.1 Selection of ANN Model Rainfall Inputs

As described in Chapter 3, the appropriate choice of ANN inputs is important to ensure good model performance. Therefore, an initial analysis was conducted to establish the appropriate rainfall inputs for the model. The nine 1 km^2 OS radar rainfall grid squares in proximity to the CSO chamber are displayed in Figure 5-6. Four OS grids were identified as containing upstream sewer pipes directly contributing to the CSO. To determine if the selection of the rainfall input grid

Table 5-2 Summary table of the characteristics of the CSO case study sites utilised in the ANN model preliminary tests.

CSO	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Location	Wirral	Cumbria	Cumbria	Cumbria	Lancashire	Lancashire	Carlisle	Greater Manchester	Cumbria	Greater Manchester
Catchment Type	Urban	Urban	Rural	Urban	Urban	Urban	Urban	Urban	Urban	Urban
Upstream Network Size	Medium	Large	Small	Small	Medium	Small	Large	Small	Small	Large
CSO Chamber Height	1700	4350	1560	1500	2110	2000	2670	2880	2210	2200
CSO Spill Level (mm)	520	210	390	150	310	200	380	1320	470	840
Training Datasets	22/05/16 - 27/12/16	13/07/16 - 06/01/17	13/07/16 - 06/01/17	20/03/16 - 03/07/16	06/07/16 - 01/02/17	06/09/16 - 07/02/17	15/11/16 - 28/02/17	11/03/16 - 16/12/16	7/8/16 – 4/1/17	9/9/16 - 7/1/17
Testing Datasets	28/12/16 - 31/03/17	07/01/17 - 22/03/17	06/01/17 - 22/03/17	03/07/16 - 16/08/16	01/02/17 - 01/05/17	07/02/17 - 13/04/17	28/02/17 - 13/04/17	16/12/16 - 14/04/17	5/1/17 – 14/4/17	8/1/17 – 28/3/17

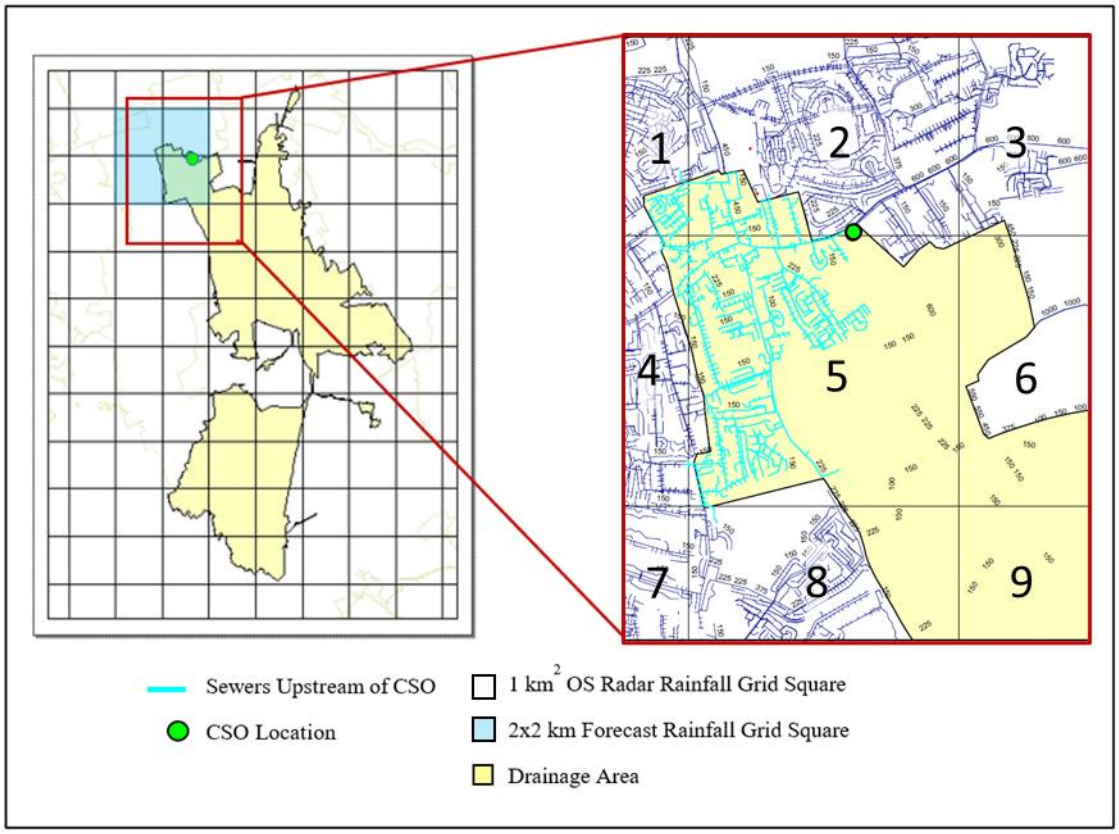


Figure 5-6 Case study CSO location showing forecast rainfall and radar rainfall grid squares.

square affects the performance of the ANN to the model the following tests were conducted.

First, an analysis was performed to determine which rainfall grid squares (sub-catchment areas) contribute most significantly to the CSO level. This was accomplished by calculating the cross correlation between the rainfall data and the level. The cross correlation is a time-series analysis technique used to measure the similarity between two signals in relation to their time lag. The maximum of the cross-correlation function indicates the time lag where the signals are best aligned. The cross correlation is frequently used in hydrology to understand in relationship between precipitation and ground-water levels (McCoy & Blanchard, 2008; Tirogo et al., 2016) and has also been used previously to determine the relationships between CSO level and rainfall grid squares when designing ANN models (e.g. Fernando et al., 2006; Kurth et al., 2008; Mounce et al., 2014b). Further information regarding the cross correlation is given in Section 4.3.1

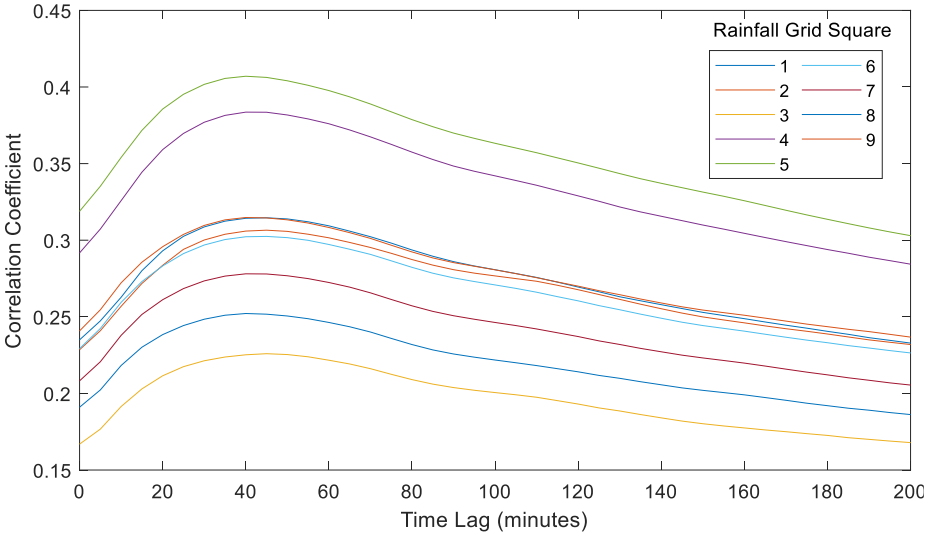


Figure 5-7 Cross correlations between measured CSO level data and rainfall intensity data from surrounding OS grid squares.

The cross correlation also gives a good indication of the most suitable number of past time steps to use as inputs to the ANN model. The delay between a rainfall event and the time when the runoff impacts the CSO level should be considered when selecting the size of past rainfall data to feed the model.

Figure 5-7 presents the cross correlations calculated for a range of time lags for each of the identified rainfall radar cells, over the training dataset. As can be seen, the cross-correlation of each rainfall grid square demonstrates similar behaviour - increasing with increasing lag time, reaching a maximum at approximately 45 minutes, before decreasing. Forty-five minutes is close to an estimated time of concentration for the catchment - the time needed for water to flow from the most remote point in a watershed to the point of collection.

As expected, grid square 5, which contains a large number of sewer pipes upstream of the CSO chamber, produces the highest cross-correlation. An examination of the other 9 CSO sites analysed, indicates that generally rainfall data from the OS grid square in which the CSO chamber is situated produces the highest cross correlation value.

Next, an analysis was conducted to examine if utilising data from different CSO grids effects the EANN model performance. ANN models were trained using rainfall data from each of the 9 rainfall grid squares. In addition, a model was

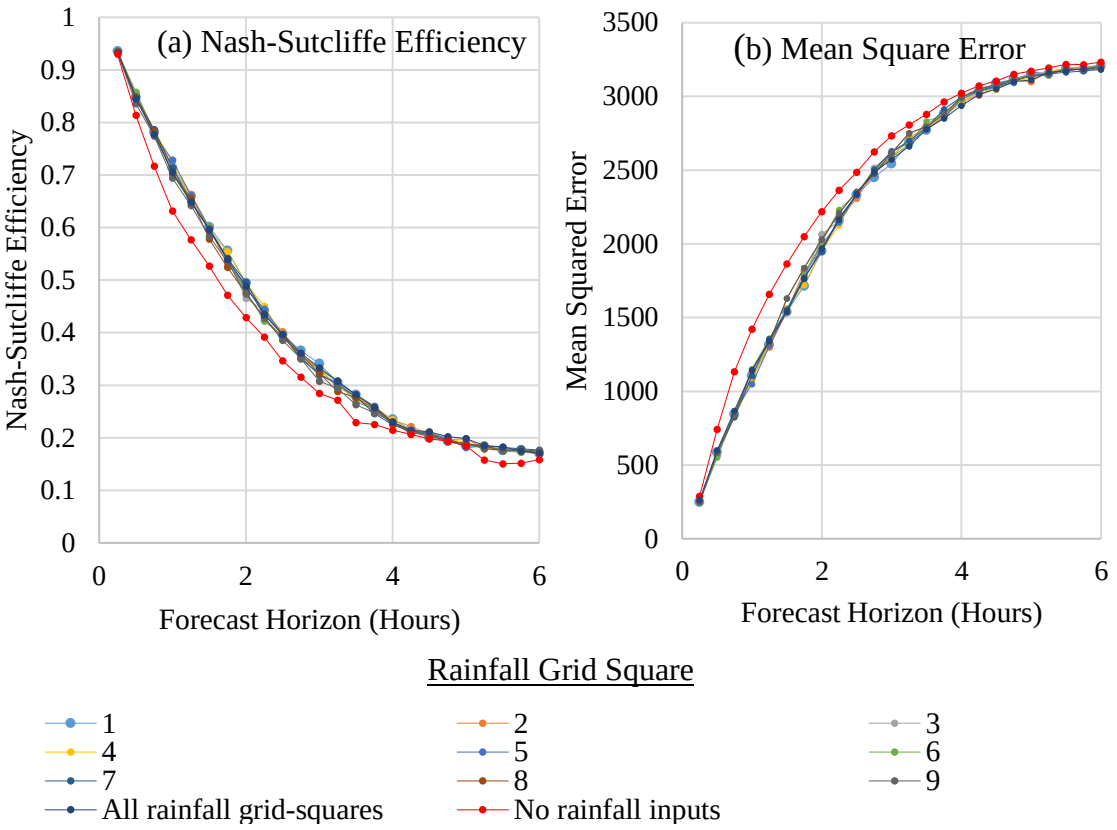


Figure 5-8 Comparison of the (a) Nash-Sutcliff efficiency, and (b) the mean squared error, of ANN CSO level forecasts vs measured CSO level when utilising rainfall data from various rainfall grid squares.

constructed utilising data from all 9 rainfall grid squares, to determine if utilising data from multiple sources can improve the model results. A model was also constructed using no rainfall inputs to analyse the benefit of using rainfall as an input.

The models were trained and tested using the datasets described in Table 5-2, for forecast horizons from 15 minutes to 6 hours ahead at 15-minute intervals (i.e. 24 forecast horizons in total). Figure 5-8 displays the Nash-Sutcliff efficiency and the mean squared error of the ANN model forecasts. As can be seen, the choice of rainfall grid appears to have virtually no effect on the model results. This is surprising as the cross-correlation presented in Figure 5-7 indicates that certain grid squares (generally those containing sewer pipes upstream of the CSO chamber) have a greater relation to the CSO level. It was assumed therefore that using data from these grids would produce superior results. As expected, utilising no rainfall data produces inferior results – indicating that useful information is

inherent in the rainfall data. Similar results were observed in the all the additional CSO sites analysed.

The similarity of the rainfall data from the various grid squares was analysed using the Kendall rank correlation coefficient (also known as Kendall's τ coefficient (Kendall, 1975). The Kendall rank coefficient is a rank-based non-parametric (i.e. distribution-free) statistical test designed to measure the strength of the association between two variables. It is widely used to assess precipitation data (e.g. Sheng Yue et al., 2002, Tirogo et al., 2016, Zhao et al., 2015). The Kendall coefficient was applied here, rather than the R^2 coefficient for example, as it does not require the assumption that the relationship between variables is linear or normally distributed (Yue et al., 2002) (as the R^2 coefficient does). Under the null hypothesis, i.e. when there is no relationship between the two variables, the Kendall coefficient has an expected value of zero, while a value of 1 indicates a perfect relationship.

The Kendall coefficient of the rainfall data from each OS grid square, when compared to the data from grid 5 (the square containing the CSO chamber) is displayed in Table 5-3. The values are all greater than 0.89, indicating a high degree of similarity between the rainfall datasets. Therefore, it is possible that the selection of rainfall grid ultimately does not have a significant effect on the ANN model performance as the data is very similar. Similarly, including data from additional rainfall grids does not result in improved results as little extra information is being provided.

Table 5-3 Mann Kendall coefficient of rainfall data from each OS grid square compared to grid square 5.

Grid Square	1	2	3	4	6	7	8	9
Kendall Coefficient	0.898	0.919	0.908	0.921	0.919	0.908	0.920	0.901

Therefore, based on this analysis it was determined that utilising rainfall data from the OS grid square containing the CSO chamber only is sufficient, and that including rainfall inputs from additional squares does not appear to measurably improve the ANN model results. This is advantageous as it is therefore not necessary to calculate the cross-correlation or use the upstream sewer network information to determine the appropriate grid square for each new site the ANN is applied to, therefore reducing the setup required.

5.5.2 Selection of ANN Model Inputs

Time series rainfall and CSO level data is fed to the ANN model as a sliding window of lagged level and rainfall inputs. To optimise model performance, the ANN model performance was evaluated for different sizes of input windows to

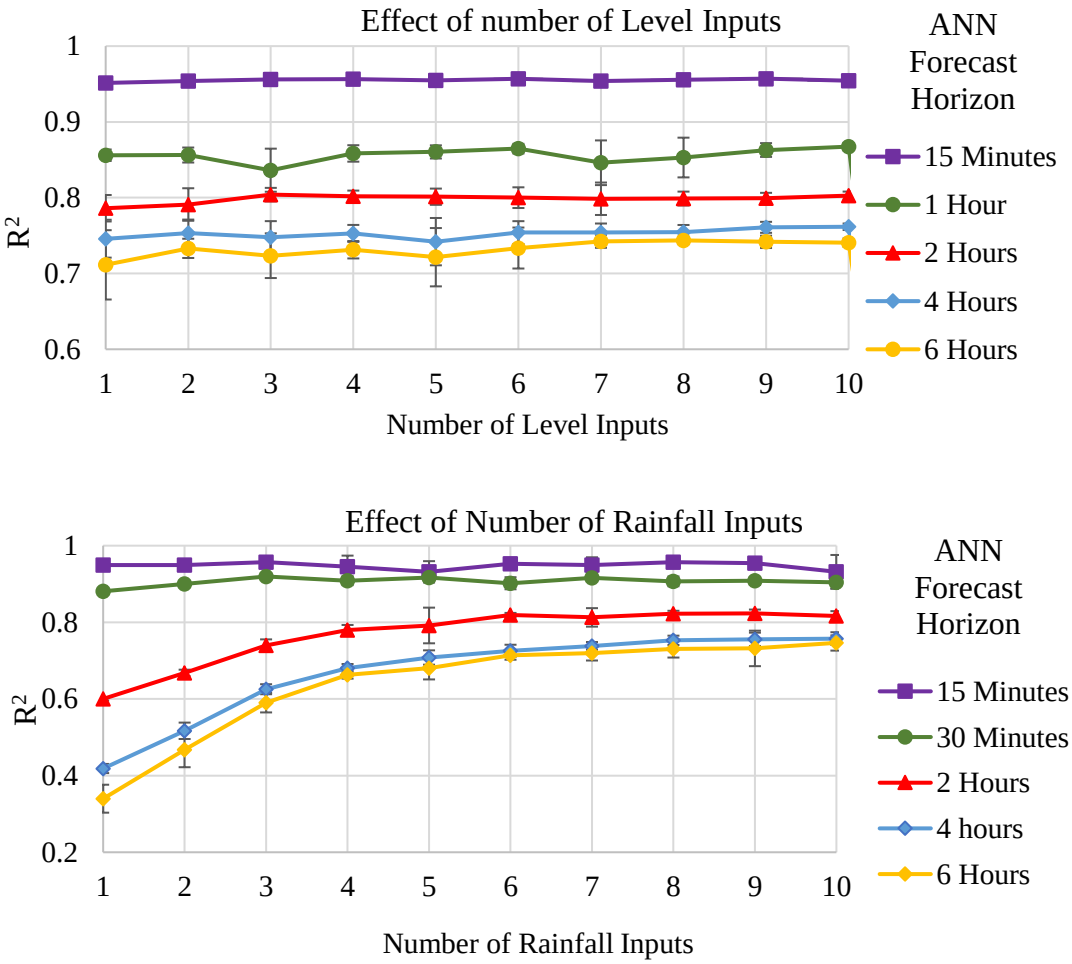


Figure 5-9 R^2 of ANN forecast CSO Level vs measured level for different numbers of level and rainfall inputs when utilising perfect forecast rainfall data.

determine the best performing value. This was performed independently for rain and level input windows. Window sizes from 1 to 10 timesteps were analysed, with a timestep length of 15 minutes. For each window size ANN models with forecast horizons from 15 minutes to 6 hours, at 15-minute intervals, were constructed (i.e. 24 models in total). These ANNs were run 10 times and the average R^2 value used to assess the model performance, with the standard deviation shown as error bars. The results are presented in Figure 5-9.

Varying the number of level inputs is demonstrated to have a minimal effect on the model performance. Similarly, varying the number of rainfall inputs at low forecast times does not significantly affect the performance. However, when forecasting 2 or more hours ahead, increasing the number of rainfall inputs considerably improves the models' accuracy. For example, the R^2 value forecasting 6 hours ahead increases from 0.34 (1 rainfall input) to 0.75 (10 rainfall inputs).

5.5.3 Evaluation of the Inclusion of Forecast Rainfall Data

Rainfall nowcasts have improved considerably since they were first developed. Nowcasts have been demonstrated to significantly increase the accuracy of hydrological forecasting e.g. for flood models (Berenguer et al., 2005; Sharif et al., 2006; Zanchetta & Coulibaly, 2020) and urban control (Thorndahl & Rasmussen, 2013). However, nowcasts still suffer from sources of inaccuracy, especially when forecasting at high lead times. Their predictive skill is affected by both the errors affecting the radar rainfall measurements and the limitations of

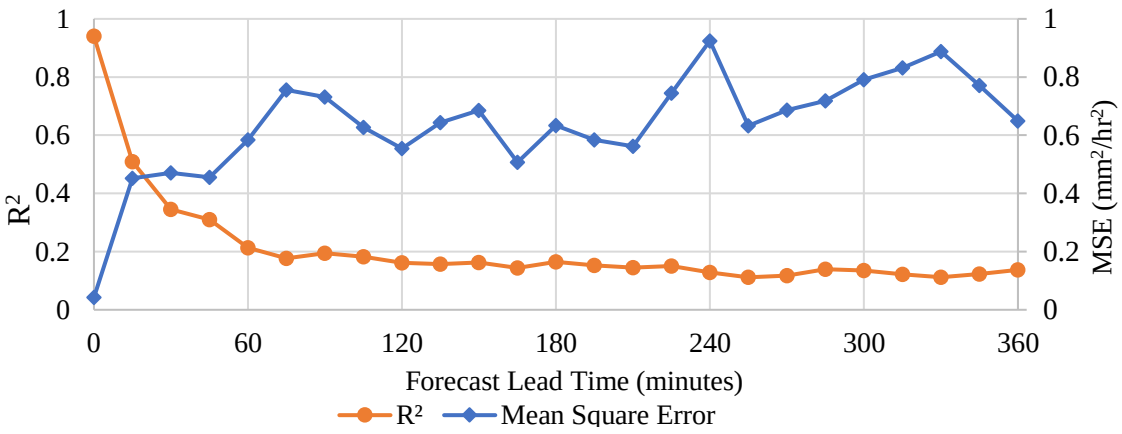


Figure 5-10 R^2 and MSE of forecast rainfall data compared to radar rainfall data as a function of forecast lead time.

the models designed to predict rainfall motion and evolution (Liu et al., 2019). A number of analyses of the MET office nowcast data have demonstrated this decrease in accuracy, e.g. Clark et al., 2016; Golding et al., 2014; Simonin et al., 2017. This is illustrated in Figure 5-10 (copied from Figure 3-5) which compares the accuracy of MET office rainfall nowcast data against observed (i.e. actual) radar rainfall data for the rainfall grid square of the CSO catchment described above. As can be seen, the forecasts are in good agreement at low lead times. However, the forecast rainfall accuracy decreases rapidly for higher forecast lead times greater than 30 minutes. It was therefore hypothesised that using highly inaccurate data as an input to the ANN model could be detrimental.

Thus, to evaluate the benefits which can be gained by using rainfall nowcasts, and to understand the knock-on effects of the inaccuracies in the data, three different versions of the ANN model were assessed and compared.

i. No nowcast data ANN: this model uses inputs of antecedent observed rainfall data and CSO level data only.
ii. Perfect nowcast data ANN: in addition to the observed rainfall and level data this model uses perfect rainfall forecasts (i.e. forecasts assuming perfect knowledge of historical rainfall into the future). This model was analysed to understand the effect that forecast inaccuracies have on model performance and to investigate the maximum level of prediction improvements that can be obtained using rainfall forecast data.
iii. Actual nowcast data ANN: The third model uses actual forecast rainfall data (i.e. real forecasts obtained from the UK Met Office) in addition to observed rainfall data and level data.

The three ANN-based models were applied CSO site case study 1, described above. Predictions for future CSO chamber depth were performed from 15 minutes to 6 hours ahead at 15-minute intervals. All results shown are from the test, i.e. unseen, dataset.

Aside from the inclusion/exclusion of rainfall nowcasts data the structure of the 3 ANN models are identical. Based on the analysis described in Section 0 the number of level inputs was set as 5, and the number of rainfall inputs varied depending on the forecast lead time being predicted. Regarding the construction of the ANN architecture for the three models, no universally accepted general

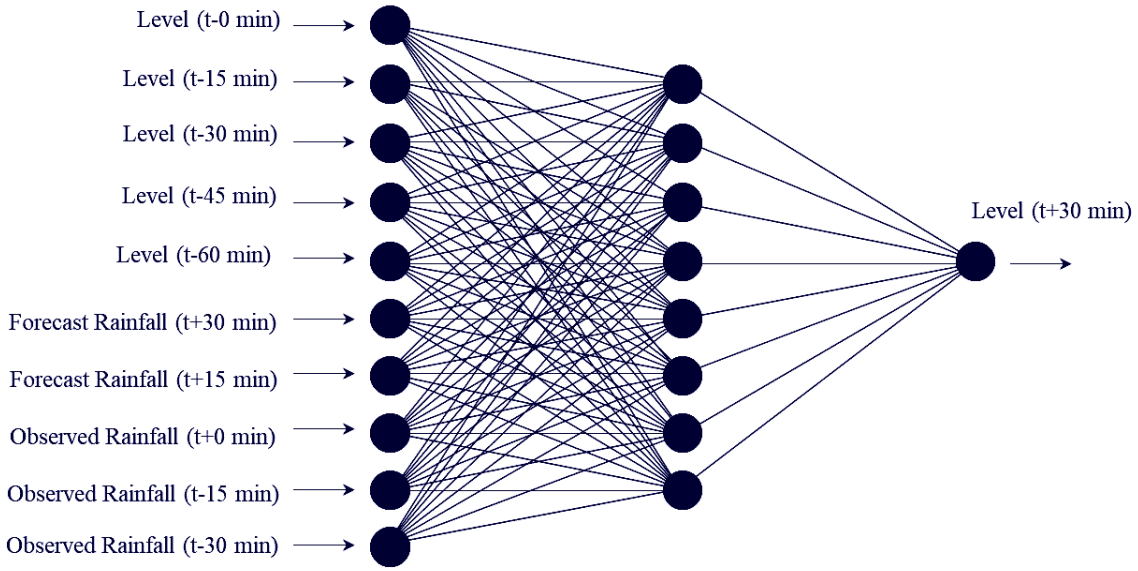


Figure 5-11 ANN architecture predicting 30 minutes ahead with actual forecast data.

rule for calculating the optimum number of neurons in the hidden layer for an ANN exists, though the common recommendation is $N_{HiddenNeurons} = \frac{2}{3}(N_{Inputs} + N_{Outputs})$. An initial sensitivity study (not shown here) indicated that an ANN model with 8 neurons produced the best results, which is generally in line with that rule of thumb. A representation of the resulting architecture for ANN models (ii) and (iii) when predicting 30 minutes ahead is illustrated in Figure 5-11. The model contains 5 rainfall inputs: a combination of forecast rainfall data (15-30 minutes future forecasts) and past observed radar rainfall data (0 – 30 minutes past).

The Nash-Sutcliffe efficiency and mean squared error values of the ANN models when modelling different forecast times are presented in Figure 5-12. The models were run 10 times for each forecast horizon, with the standard deviation of the results presented as error bars.

The results demonstrate that including forecast rainfall significantly improves the ANN model performance, most significantly at high forecast lead times. All three ANN models successfully produce good forecasts up to 1 hour ahead, with Nash-Sutcliffe efficiency values greater than 0.7. As expected, the accuracy of the models decrease as the prediction horizon increases. However, the inclusion of both perfect and actual forecast rainfall data enables the relevant ANN model to predict the CSO level values farther into the future with greater accuracy. The

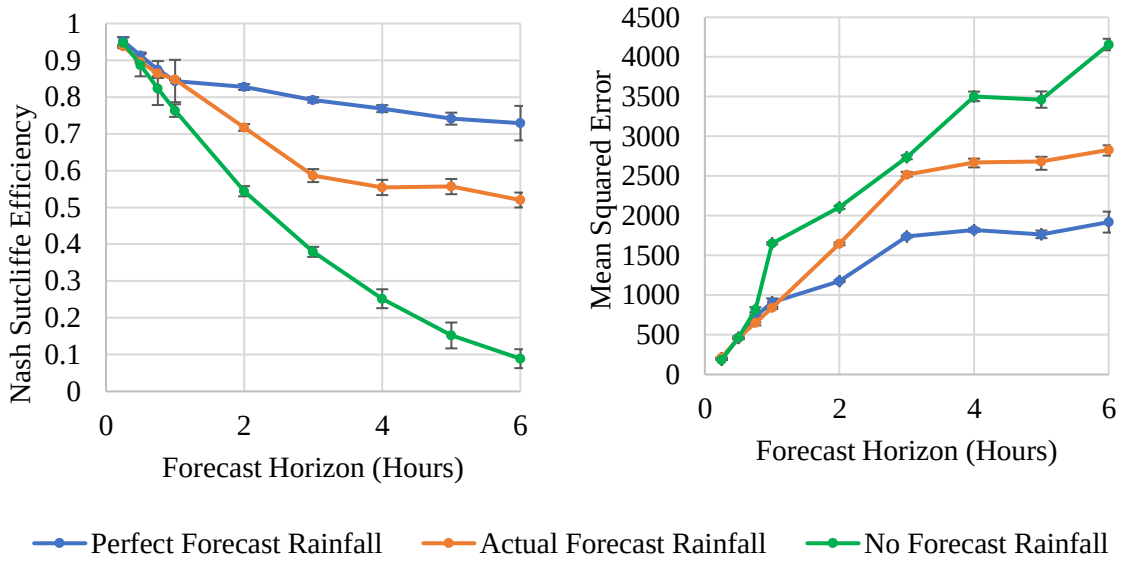


Figure 5-12 Comparison of ANN model performance with (i) perfect forecast rainfall, (ii)actual forecast rainfall and (iii) no forecast rainfall data.

perfect forecast model produces better results at high forecast horizons, indicating that the inaccuracies in the nowcast data in turn affect the accuracy of the ANN model predictions. However, the performance of the ANN model using actual forecast data is significantly superior to the model with no forecast data.

Figure 5-13 compares CSO level predictions 15 minutes and 120 minutes ahead for a time period containing three sets of spill events corresponding to periods of increased rainfall. The 15 minutes ahead prediction shows that all three ANN model predictions demonstrate almost perfect agreement with the measured level data (demonstrated by the R^2 values greater than 0.9). The four spill events are all correctly forecasts – with the timings and magnitude of the overflows accurately forecast. This is representative of the entirety of the test dataset – 100% of the spill events are correctly forecast.

When predicting 120 minutes into the future the prediction accuracy noticeably deteriorates, evident also in the decreased R^2 values. The ANN model without forecast data is unable to predict any of the overflow events and significantly under-forecasts the increases in level due to rainfall. The inclusion of perfect and actual forecast rainfall enables the relevant ANN model to capture the relationship between CSO level and rainfall with greater accuracy, increasing the R^2 value from 0.54 (no forecast) to 0.74 with actual forecast data and 0.83 with perfect forecast data. The actual forecast model is able to generate a warning for the last

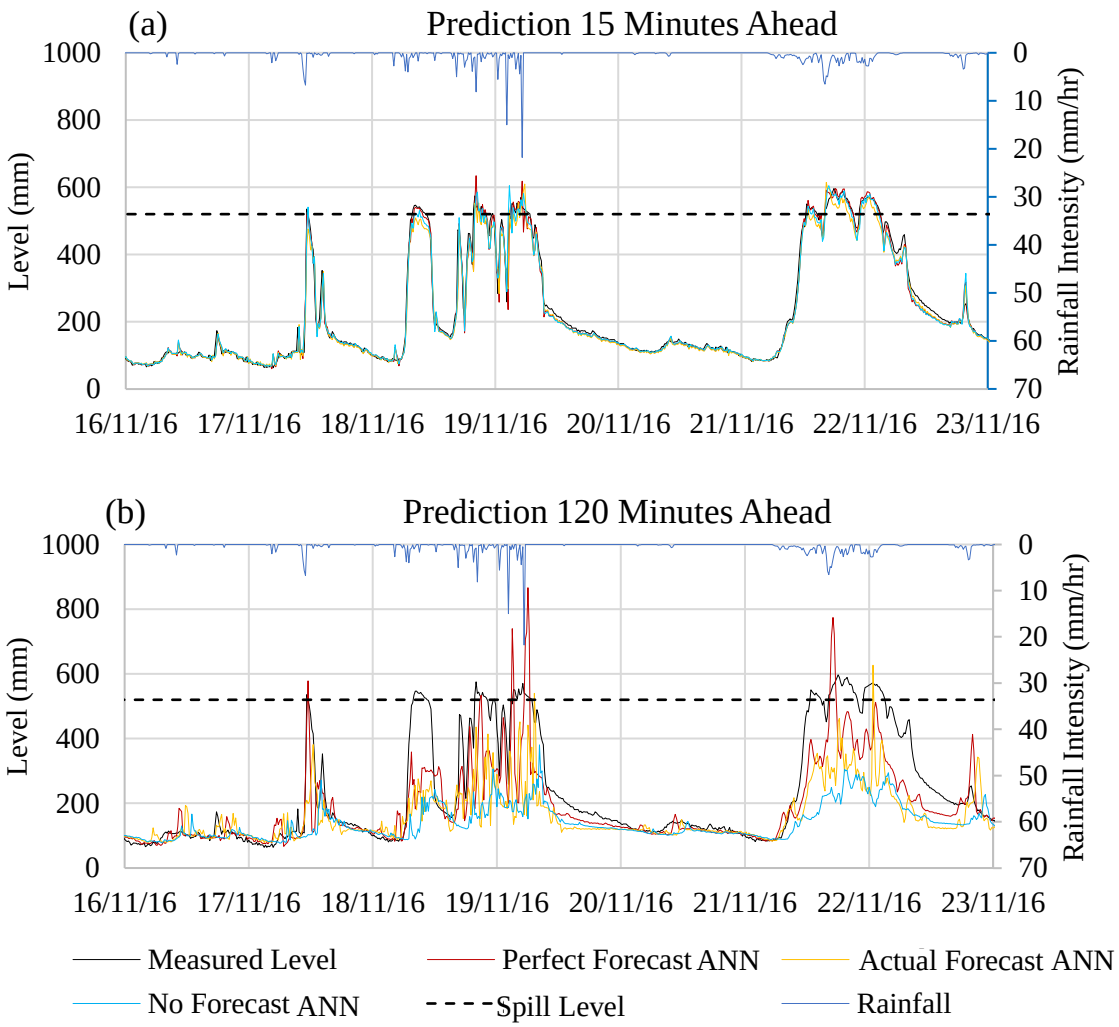


Figure 5-13 Modelled CSO level prediction vs measured level (a) 15 minutes and (b) 120 minutes ahead for selected time period.

2 spills while the perfect forecast model correctly predicts overflows for all 3 sets of spills.

It is evident therefore that the use of forecast rainfall as an input measurably improves the accuracy of the ANN model forecasts and increases the prediction range. The decrease in the accuracy of the forecast rainfall appears to affect the accuracy of the ANN CSO predictions to an extent. However, even at these large forecasts horizons the ANN model using actual forecast rainfall data is significantly better than the model with no forecast rainfall data.

5.5.4 Section Summary

A series of preliminary analyses have been presented here, which were used to determine certain initial design parameters for the ANN CSO level prediction model.

Firstly, it has been demonstrated that selecting rainfall data from the OS grid square containing the CSO chamber is sufficient – it is not necessary to perform an initial analysis to first determine which grid squares contribute most significantly to the CSO level. This is beneficial as it reduces the initial set-up required before building the model.

Secondly, based on a limited analysis conducted here it seems that the inclusion of forecast nowcast data substantially improves the ANN model prediction performance. It was anticipated that the inaccuracies in the nowcast data at high forecast horizons could be detrimental to the model results, however, this was demonstrated to not be the case. Indeed, the improvements gained by using nowcast data is most substantial at high lead times. Previous ANN CSO level prediction models (e.g. Mounce et al., 2014a; Mounce et al., 2014b; Zhang et al., 2018b) have not utilised nowcast data, so it is valuable to demonstrate that they can be used to significantly improve model forecasts.

The analyses presented here were performed using an ANN with inputs determined via trial and error. Further analysis of these effects using a CEANN model have not been performed. However, it is assumed that the same benefits/results will apply. Therefore, based on these tests, all ANN models presented hereafter (i) include rainfall nowcast data as an input, and (ii) utilise only rainfall data obtained from the OS grid square containing the CSO chamber.

5.6 Testing of the CEANN Model

In the analysis described in this section the CEANN model presented in Chapter 3 has been applied to four case study sites, in order to test its prediction capabilities. As described in Chapter 3, in order to fully evaluate the performance of the CEANN model, and understand the benefits, it has been compared to three other ANNs:

(i) An ANN model with structure and parameter determined via trial and error (representing CSO modelling approaches currently adopted in the literature),
(ii) An All-weather EANN, i.e. a single EANN model applied simultaneously to both wet and dry weather,
(iii) A Combined Wet/Dry weather EANN, i.e. a model utilising wet and dry EANNs combined using a hard switch between the models rather than a weighted average.

All four model types were applied to the four real case study sites in the United Utilities network. Information relating to the sites is displayed in Table 5-4. The sites were selected to represent the characteristics of different CSOs in the UK; the upstream catchments and sewer networks are of different sizes and are situated in both rural and urban regions and the CSO chambers are also of various sizes.

The time periods of the data used for training, validate and testing the models varies for each site, due to differences in the level monitors' installation date, and in some cases due to malfunctions and recalibration of the monitors. However, all sites were required to have a minimum of 8 months of training data. Tests indicated that the model performance did improve when the size of the training datasets were increased. However, a minimum training dataset of three months was still demonstrated to produce good results. For all sites 50% of data was used for model training and 25% was used to validate the trained models. The remaining 25% of data was used to test the model in a simulated online fashion, i.e. feeding the incoming data to the models as they would operate in real-time, to ensure that the results obtained are representative of what will produced when utilised by a utility real-time. All the model results shown here are from these test datasets.

Table 5-4 Summary table of CSO case study sites' characteristics used to assess the CEANN model performance.

CSO ID	1	2	3	4
Location	Carlisle	Cumbria	Wirral	Wirral
Catchment Type	Urban	Rural	Urban	Urban
Upstream Network Size	Small	Medium	Medium	Large
CSO Chamber Height	1790	2235	1700	1560
Spill Level (mm)	1090	1000	520	960
Average Level (mm)	187	147	102	147
Training Dataset	1/3/2016 – 30/12/2016	1/6/2016 – 30/4/2017	18/4/2016 – 10/3/2017	1/3/2016 – 30/1/2017
Validation Dataset	31/12/2016 – 31/5/2017	1/5/2016 – 14/10/2017	11/32017 – 20/8/2017	31/11/2017 – 17/7/2017
Testing Dataset	1/6/2017 - 30/10/2017	15/10/2017 - 30/3/2018	21/8/2017 - 30/1/2018	18/7/2017 - 1/1/2018

5.6.1 ANN Model Parameter Selection via Trial and Error

Model (i), the conventional ANN, has an input structure and architecture determined manually via trial and error for each CSO site it is applied to. This process is performed independently for each parameter and is accomplished by testing and training model the 24 ANN models (from 15 minutes to 6 hour lead times) with varying values and selecting the value which produces the best results generally over the 24 forecast horizons (evaluated using the Nash-Sutcliffe efficiency). The decision variables, and their ranges of variability, are presented in Table 5-5 (copied from Table 3-1).

The graphs constructed to select the best performing parameter values for case study site 1 are presented in Figure 5-14. For each parameter a minimum of three different values were considered and the best performing value selected. It is probable that better results could be produced by selecting the parameter values separately for each of the 24 forecast horizon models. However, this would be significantly more time consuming and labour intensive. It should also be noted that this trial and error method does not consider the interaction between different parameters – however, again, incorporating this would be extremely time consuming.

Table 5-5 Decision variables and associated ranges of variability of the ANN model.

Decision variables	**Range of values used in optimisation**
Level data lag size (number of time steps)	2 – 72
Radar rainfall data lag size (number of time steps)	2 – 72
Forecast rainfall data lag size (number of time steps)	1 - 24
Time of day	Use/ Do not use
Day of week	Use/ Do not use
Number of hidden neurons	10 - 100
Number of training cycles	50 – 500
Coefficient of weight decay regularisation (α)	0 - 1000

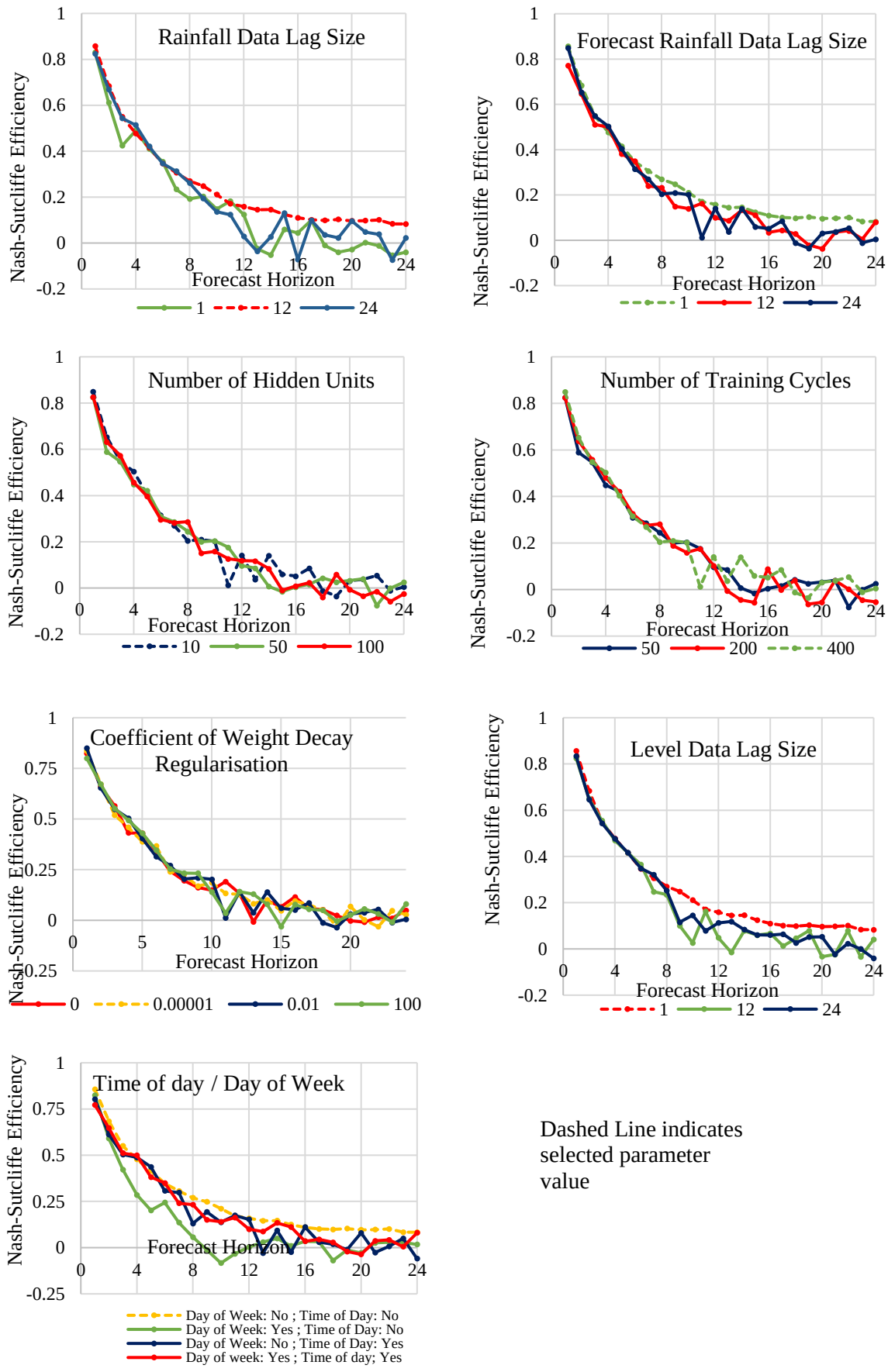


Figure 5-14 Assessment of various ANN parameters values for case study site 1, used to select the best performing values.

5.6.2 EANN Model Parameter Selection

The parameters for the CEANN and EANN models are selected by an Evolutionary Algorithm, which automatically selects the optimal (i.e. that yields the best forecasting performance) ANN input structure and parameter set for the specific training data and forecast horizon considered. The forecasting performance is here measured using the Nash-Sutcliffe efficiency.

Table 5-6 presents the parameters selected by the EA for the All-weather EANN model (i.e. model (ii)), and the dry weather model, and the wet weather model (utilised by the CEANN model and model (iii)). The results are presented for selected forecast horizons for CSO case study site 1 only due to space limitations. From Table 5-6 it appears that the selection of parameters by the EA is influenced noticeably by the forecast horizon and model type. Some trends in the parameter selection can be observed. For example, the number of antecedent rainfall and forecast rainfall inputs generally increase for higher forecast horizons for all the models, this is presumably because these horizons are more difficult to predict and so more data is beneficial. The time of the day is not utilised by the wet weather model - during wet weather the CSO level is influenced predominantly by the rainfall and so including the time of day is not beneficial. During dry weather, however, CSO level is highly correlated to the time of day and so this parameter is generally used by the dry weather and All-weather model at forecast horizons greater than one hour. In contrast the day of the week is utilised only by the wet weather model during high forecast horizons – this may be because higher lead times are more difficult to predict, therefore additional information is valuable in producing more accurate results. For some parameters no clear pattern is evident, namely the selection of the number of hidden neurons, the number of training cycles or the coefficient of weight decay regularisation.

A similar pattern in the parameters selected by the EA was observed for all the case study sites. However, the actual values selected for each model varied considerably. Therefore, using a predetermined set of parameters based on these results, rather than using an EA to evolve the optimal parameters for a given location, will likely produce inferior results.

Table 5-6 Automatically selected (i.e. optimal) EANN parameters & input structure for all weather, dry weather, and wet weather models for selected forecast horizons.

	All Weather EANN								Dry Weather EANN								Wet Weather EANN							
Forecast Horizon	Day Of Week	Time Of Day	Level Data Lag Size	Rain Data Lag Size	Forecast Rain Data	No of Hidden Units	No of Training	Alpha	Day Of Week	Time Of Day	Level Data Lag Size	Rain Data Lag Size	Forecast Rain Data	No of Hidden Units	No of Training	Alpha	Day Of Week	Time Of Day	Level Data Lag Size	Rain Data Lag Size	Forecast Rain Data	No of Hidden Units	No of Training	Alpha
0.25	no	no	6	6	2	60	350	1	yes	no	4	4	22	10	500	0.001	no	no	4	18	1	90	200	0.01
0.5	no	no	18	4	2	100	300	0	no	no	12	4	4	70	350	0.1	no	no	4	12	2	60	400	0.1
0.75	no	no	18	4	8	80	450	0	no	no	6	4	4	60	400	0.001	no	no	6	12	2	100	350	1
1	no	no	12	12	10	60	350	0	no	no	6	6	4	40	500	1	no	no	4	12	4	30	150	0.00001
1.25	no	no	12	6	10	70	300	0.1	no	yes	12	24	22	20	400	0.01	no	no	30	30	4	80	150	0.01
1.5	no	no	18	12	14	100	350	0.1	no	yes	4	18	18	20	400	0.1	no	no	30	12	14	20	50	0.00001
2	no	yes	18	48	22	80	300	10	no	yes	6	6	10	50	500	0.001	no	no	4	4	1	90	400	1
2.25	no	no	6	12	14	60	300	0.0001	no	yes	18	48	14	80	250	0.001	no	no	6	24	24	100	50	1
2.5	no	yes	6	72	24	90	450	10	no	yes	4	18	14	10	200	0	no	yes	24	12	4	40	50	10
3	no	yes	6	6	24	10	200	0.0001	yes	yes	4	48	14	30	500	0.0001	no	no	18	42	18	10	250	1
3.25	no	yes	18	30	18	80	500	0.0001	yes	yes	6	4	24	50	400	0	yes	no	24	30	24	40	400	10
3.5	no	yes	18	42	22	40	400	0.00001	no	yes	42	48	22	30	350	0.0001	no	no	6	48	24	100	50	10
4	no	yes	18	30	24	70	300	0	no	yes	12	6	18	10	400	0.01	yes	no	24	24	22	30	500	10
4.25	no	yes	12	36	22	10	500	0.1	no	yes	18	30	22	50	400	1	no	no	12	6	6	30	200	0
4.5	no	yes	12	36	22	20	500	0.01	no	yes	42	42	18	40	500	0.1	yes	no	18	42	18	60	100	0.01
5	no	yes	18	24	24	10	500	0.0001	no	yes	6	18	22	40	450	0.001	yes	no	4	36	24	70	50	0.01
5.25	no	yes	12	48	24	40	500	1	no	yes	4	72	24	50	400	0.001	yes	yes	12	24	24	100	400	10
5.5	no	yes	4	18	24	40	350	0.00001	no	yes	12	42	22	20	450	0	yes	no	18	42	22	40	200	10
6	no	yes	6	42	24	40	400	0.1	no	yes	24	42	24	30	450	0.1	yes	no	18	30	8	40	50	10

5.6.3 Model Performance

This section presents the results of the application of the various ANN models to the four case study sites described above. Figure 5-15 presents the NSE, MSSIM, MSE and MAPE performance indices for CSO site 1. The indices are computed by comparing the measured level with forecast data over the entire test dataset. Each point on the graph represents a model with a different forecast horizon, from 15 minutes to 6 hours ahead. The NSE for the other three sites are presented in Figure 5-16 due to space limitations, however, it should be noted

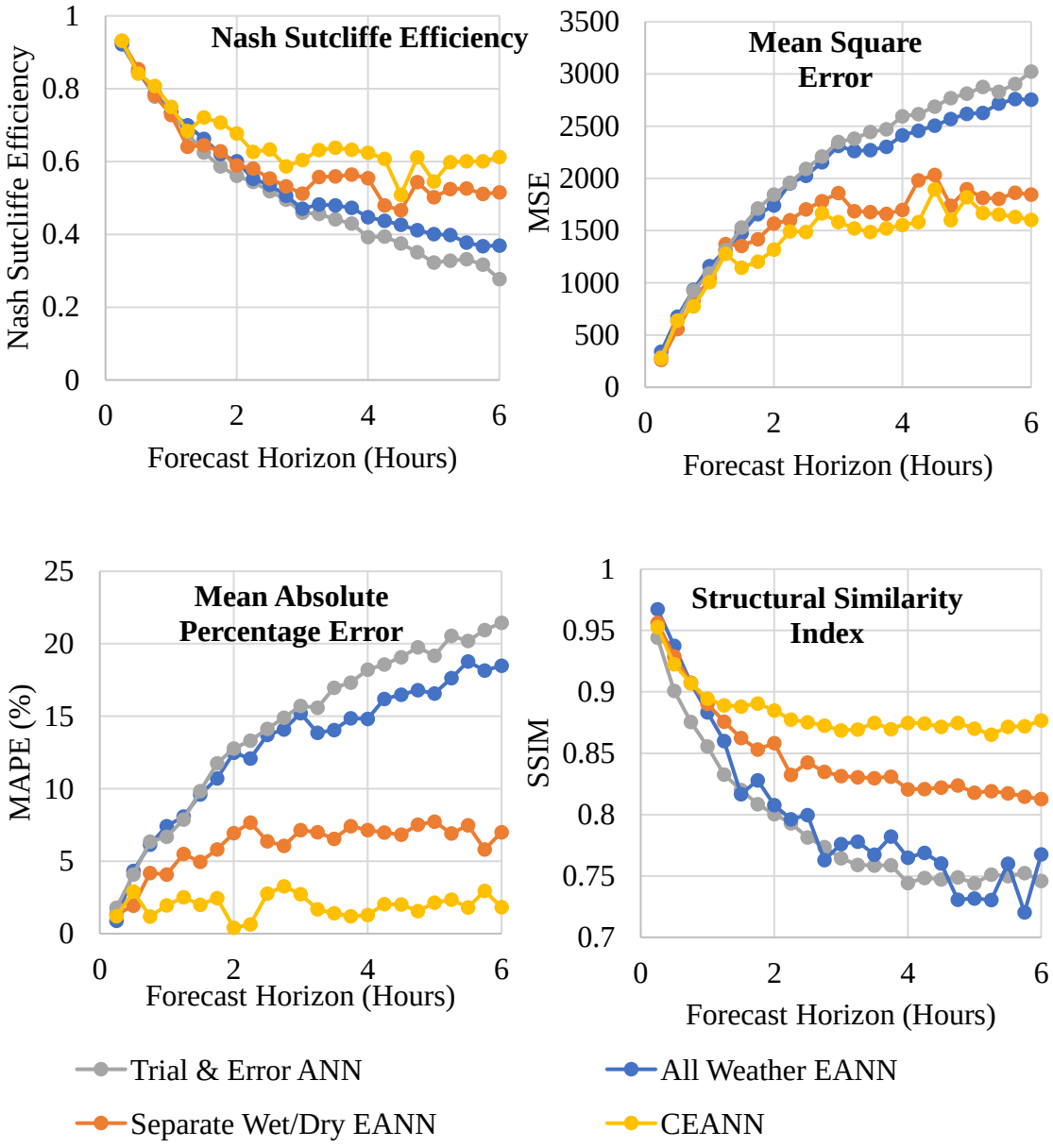


Figure 5-15 Comparison of model performance for different ANN models for CSO site 1.

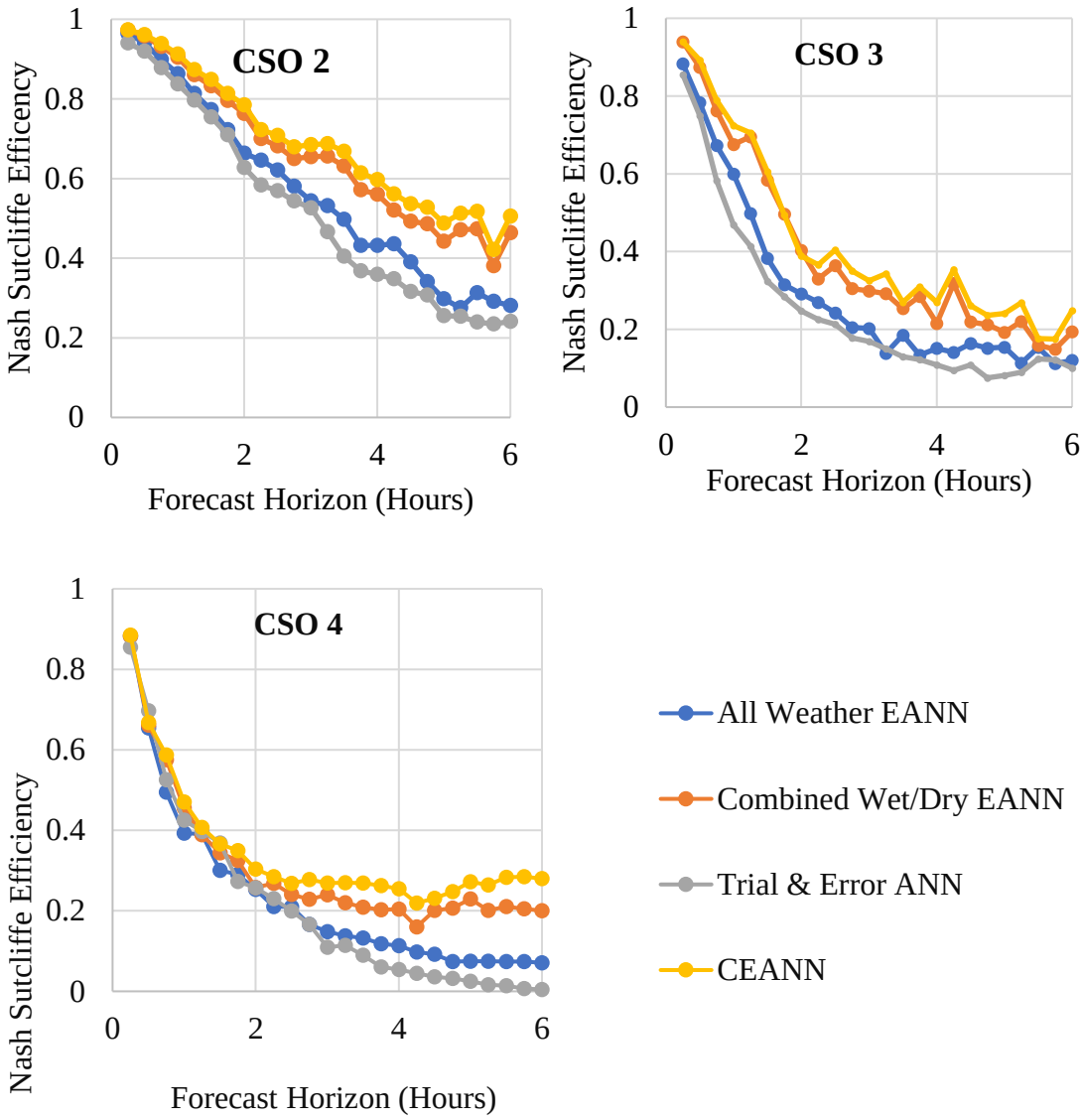


Figure 5-16 Nash-Sutcliffe efficiency for case study sites 2, 3 & 4

that each CSO case study showed similar results – factors such as catchment size and type did not appear to affect the performance of the models. All the metrics presented in Figure 5-15 and Figure 5-16 demonstrate that the CEANN model produces the best results for almost all forecasting time steps and for all CSO sites. For the few times when the CEANN does not produce the best results according to a certain metric (e.g. according to the Nash-Sutcliff efficiency at a forecast horizon of 1.25 hours the All-weather EANN produces slightly better results), the inferiority of the CEANN model is very small. Also note that this only occurs according to any performance metric not greater than two out of the twenty-four forecasting horizons and at small forecast horizons only.

In addition, the Combined Wet/Dry EANN consistently outperforms the other two models, and the All-weather EANN somewhat outperforms the Trial-and-Error ANN model. When forecasting one hour ahead both the CEANN model and Combined Wet/Dry EANN model successfully predict 67% of all spill events that occur in the validation dataset, whilst the All-weather EANN and Trial-and-Error ANN are only able to predict 42% and 33% of the spills respectively. A spill is here considered correctly predicted if the ANN model predicts the CSO level above the spill height of the chamber during the overflow event. It is evident therefore that the use of separate EANN models to overcome the issues caused by imbalanced data is effective. Furthermore, combining the models using a weighted average (i.e. the CEANN model) is shown to be superior to using a discontinuous switch approach (i.e. the Combined Wet/Dry weather EANN), although this approach still produces good results. Additionally, the EANN model is demonstrated to produce more accurate forecasts to a Trial-and-Error ANN and has the additional benefit of not requiring human input.

As can be seen in Figure 5-15 and Figure 5-16, the differences between model performances are most significant when predicting further into the future. When predicting only a small number of time steps ahead the CSO level is highly dependent on past level and so all the various ANN models are able to perform well using the most recent observed level data. At a forecast horizon of only 15 minutes all models display satisfactory results - with NSE values all greater than 0.96. However, when predicting further ahead the past and future level are less directly correlated. Thus, selecting the optimal parameters and inputs via the separate EANN methodology, i.e. tailoring the ANN model specifically to the particular forecast horizon and rainfall type, results in a greater improvement in the forecasting results. At a forecast horizon of 6 hours ahead the CEANN and Combined Wet/Dry EANN have NSE values of 0.61 and 0.52 respectively, the Trial-and-Error ANN and All-weather EANN have an NSE value under 0.4 (indicating inadequate results).

The use of separate wet and dry weather EANN models was implemented to overcome poor ANN model performance during rainfall events. These poor predictions are caused by severe data-imbalance in the rainfall dataset - wet weather timesteps are significantly underrepresented compared to dry weather timesteps. To evaluate if this technique is successful Figure 5-17 compares the

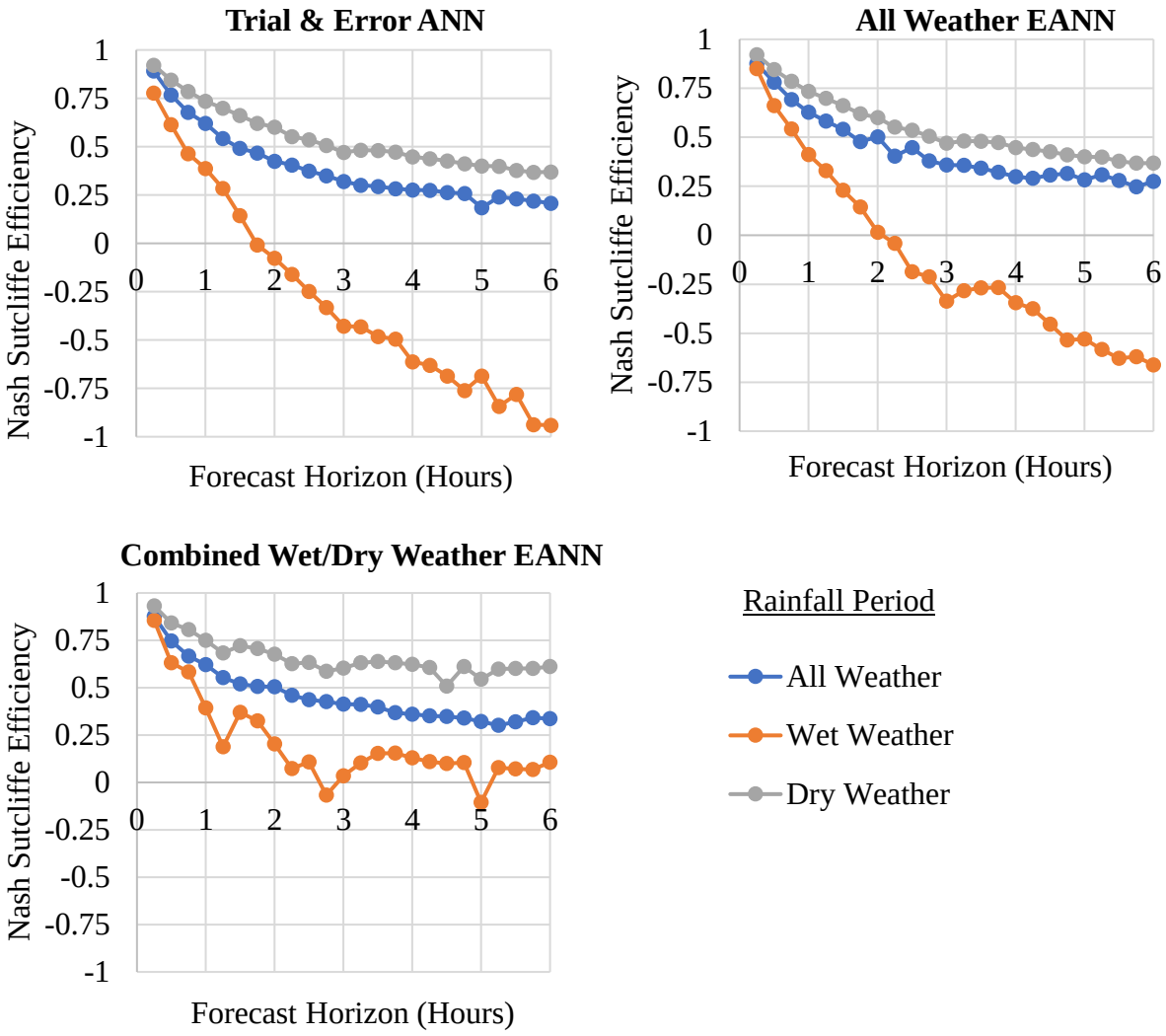


Figure 5-17 Comparison of ANN models' performance during (i) Wet weather periods, (ii) Dry weather periods and (iii) All weather periods.

performance of the Trial-and-Error ANN model (i.e. model (i)), the All-weather EANN (model (ii)) and the two EANN models optimised for wet and dry weather periods (i.e. the two individual models utilised by the CEANN model and the Combined Wet/Dry Weather EANN model). The models are compared for all weather timesteps (i.e. the entire validation period), wet weather timesteps only, and dry weather timesteps only.

The results presented in Figure 5-17 demonstrate that the separate EANN model performs slightly better than the Trial-and-Error and the All-weather EANN model for dry weather periods. However, when applied to wet weather data, the wet weather EANN model produces substantially more accurate results, especially at

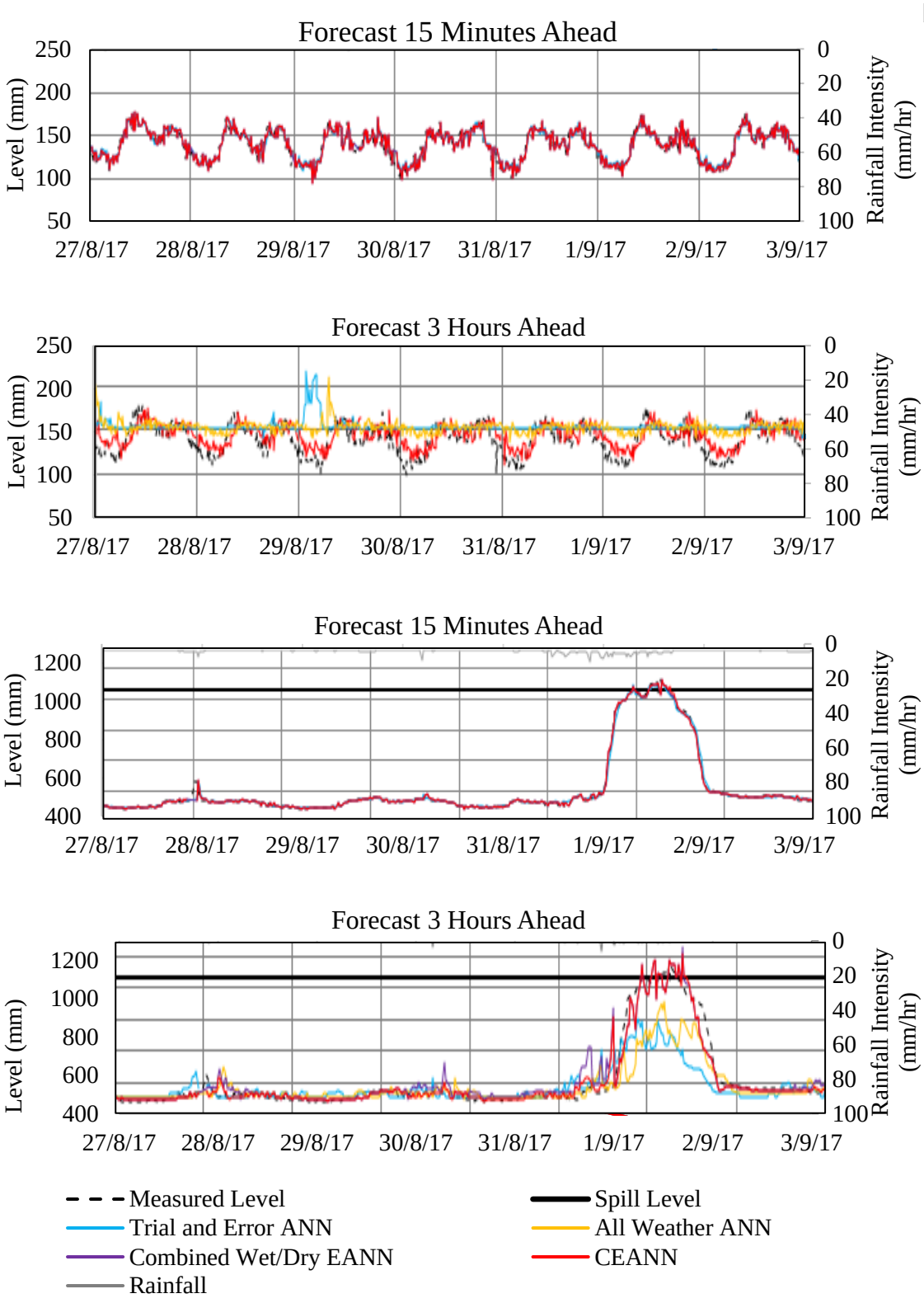


Figure 5-18 Comparison of model forecasts for different ANN models during wet and dry weather for CSO site 1.

longer forecast horizons. It appears therefore, that the use of separate wet and dry models to overcome the problem of data imbalance is successful.

Figure 5-18 displays an example of the level forecasts for each model 15 minutes and 3 hours ahead during the validation data period for CSO site 3 during: (i) dry weather and (ii) a rainfall event which causes an overflow. When forecasting only 15 minutes ahead all four ANN models accurately forecast the CSO level during both the dry and wet weather periods, precisely predicting the timings and magnitude of the spill event. When forecasting 3 hours ahead the improvements attained by optimising separate wet and dry EANN models is clearly demonstrated - during dry weather the Trial-and-Error ANN and All-weather EANN are unable to predict the diurnal dry weather level pattern, with the Trial-and-Error model forecasting an almost a straight line for all timesteps. During wet weather all the models anticipate the rise in level caused by a rainfall event, however the Trial-and-Error ANN and All-weather model significantly under predict the extent of the increase and so do not give notice of the upcoming spill event.

Analysing the entire test dataset for all the case study sites, it can be seen that the Trial-and-Error ANN and All-weather EANN models continually underestimate the CSO levels during periods of heavy rainfall and so cannot predict overflows. In contrast, the combined Wet/Dry EANN model and the CEANN model accurately forecast the dry weather and wet weather level, predicting the timing and magnitude of the spill, information that would be of great use to a wastewater utility in real time.

The 3 hour ahead CEANN prediction, and to a lesser degree the other ANN model predictions, displayed in Figure 5-18 appear smooth during dry periods, but become noisier during wet periods. This is due to the noisy nature of the rainfall data.

5.7 Overflow Prediction Performance

An import objective of the CSO level predictions is the ability to provide warnings for upcoming CSO spill events. Early detection of overflows can allow wastewater utilities implement preventative action to prevent or mitigate the effects of a spill

event. If the overflow is predicted to occur under non-compliant conditions, e.g. due to a blockage or a leak, detection is especially important as steep financial penalties could be incurred. Even if proactive management cannot be implemented to prevent the spill, the ability to forecast overflows in advance is still valuable – for example by enabling wastewater utilities to post warnings for spills which will affect bathing waters.

This section evaluates the ability of the CEANN model to correctly forecast overflows and to generate alerts in a timely and reliable manner. Two performance evaluations are conducted to analyse the model's capacity to identify the start of spill events and to correctly identify the duration and timings of the events.

It should be noted that the CEANN model has been trained to accurately forecast CSO level, rather than detect spill events, as the model has been primarily developed for use in blockage detection which requires level forecasts. Spills are a relatively infrequent event, therefore, to develop a classification neural network to detect spill events would require a very large dataset of CSO level data. The analyses have not been applied to the other ANN models as it was already demonstrated in Section 5.6.3 that they produce inferior results to the CEANN model.

5.7.1 Case Study Site

The overflow prediction analysis is applied to CSO site three, as described in A photograph of the CSO chamber and the level monitor is presented in Figure 5-19. The spill height of the chamber is 520 mm.

The first stage of the spill analysis was to identify and analyse the overflows present in the historic CSO level dataset. There is a question here concerning how to define a single spill event. The environment agency employ the 12/24 counting method, which is defined as follows (Environment Agency, 2018b): one or more overflow events within a period of 12 hours are considered to be one spill, one or more overflow events extending over a period of greater than 12 hours up to 36 hours are considered to be 2 spills. Each subsequent 24 hour

Figure 5-19 Photograph of the CSO chamber and level sensor from CSO site 1.

duration counts as one additional spill and the whole of the 24 hour block is included. However, as the CEANN model only generates forecasts up to 6 hours ahead, it would not make sense to apply this method here.

An alternate method is to count every occasion the level increases above the spill level as an individual overflow event. However, the CSO level during spills often fluctuates above and below the spill level during rainfall events (as is illustrated in Figure 5-20), and these should be defined as a one event. It was therefore decided to group spills which have less than 3 hours between the first spill end and the second spill start as one single event. Using this method 32 spills were

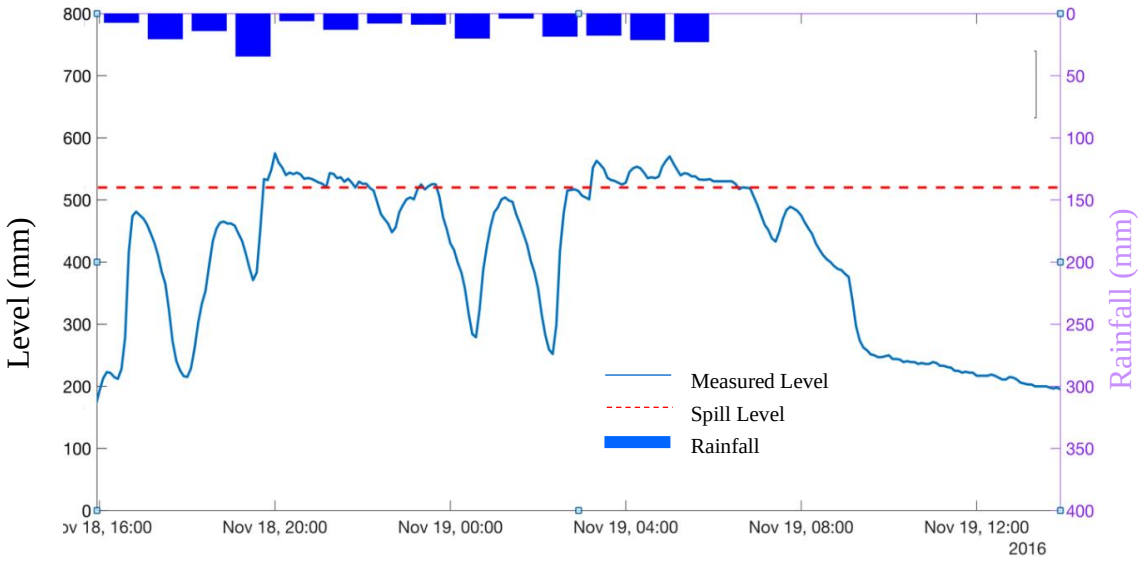


Figure 5-20 Overflow event identified in CSO level data.

identified in the analysed dataset, containing data from 14/2/18 to 1/10/18. The mean duration of these spill events is 128 minutes.

5.7.2 Overflow Prediction Analysis

When predicting spill events, several factors are important. Most simply a good model is required to successfully predict a high proportion of spills events in a timely and reliable manner (i.e. with low number of false alarms) and, ideally, with as much lead time as possible. It is also desirable that the model can predict the start time of the spill event accurately and forecast the severity of the overflow event, so that a utility can determine the most appropriate action to take.

The first performance evaluation analyses the ability of the CEANN model to correctly predict the start time of a spill event. The model performance was analysed for the 24 forecast horizons (i.e. predicting from 15 minutes to 24 hours ahead). A spill alert is generated by the model at the first time step the forecast level is above the overflow height of the CSO chamber (520 mm for this particular CSO). A correctly identified spill start is here defined as a spill alert generated by the model within 2 hours of the spill start time identified in the measured data. This tolerance period, *T1*, was allowed as the CEANN model forecasts may sometimes be a few timesteps out of sync with the measured data, however a spill detected 2 hours late is still of use to the water utility. The precise value of *T1* is not critical as small changes to the tolerance lead to small changes in the assessed performance of the model. The analysis also measures the total number of false positives produced by the model and the false alarm ratio – defined as the number of false alarms divided by the total number of forecasted events. A false alarm is here defined here as any spill alerts generated when a spill has not been identified in the measured data.

The second evaluation analyses the ability of the model to accurately predict the timing and duration of the overflow events. Correctly predicting the duration of an event is important as it is an indication of event severity. This is measured using the Jaccard similarity index, also known as Intersection over Union and the Jaccard similarity coefficient (Levandowsky & Winter, 1971). The Jaccard is a statistic designed for gauging the similarity and diversity of sample sets. It is one of the most popular similarity metrics, and is commonly applied to time series

data (e.g. Giusti & Batista, 2013; Peng et al., 2016), and to event detection data (e.g Boldt et al., 2020). The Jaccard index is defined as the size of the intersection divided by the size of the union of the analysed sample sets and is computed as follows:

$$Jaccard(A,B) = \frac{|A \cap B|}{|A \cup B|}$$

where *A* and *B* are two sets.

The Jaccard index can have a value between 0 (no temporal similarity) and 1 (perfect temporal similarity). When applied here to the CSO spills, *A* is defined as the timing of the measured spill events and *B* is defined as the timing of the forecast spill events. The index analyses the overlap between the forecast and observed spills, and so preferentially values forecasts which accurately capture both the timing and the duration of the spill events and penalises forecasts which do not – either by under or over forecasting the event length, or by predicting the spill at the wrong time.

A tolerance time, *T2*, was added of 4 timesteps (i.e. 1 hour) either side of the measured forecast start and end times. This was included as, when using 15-

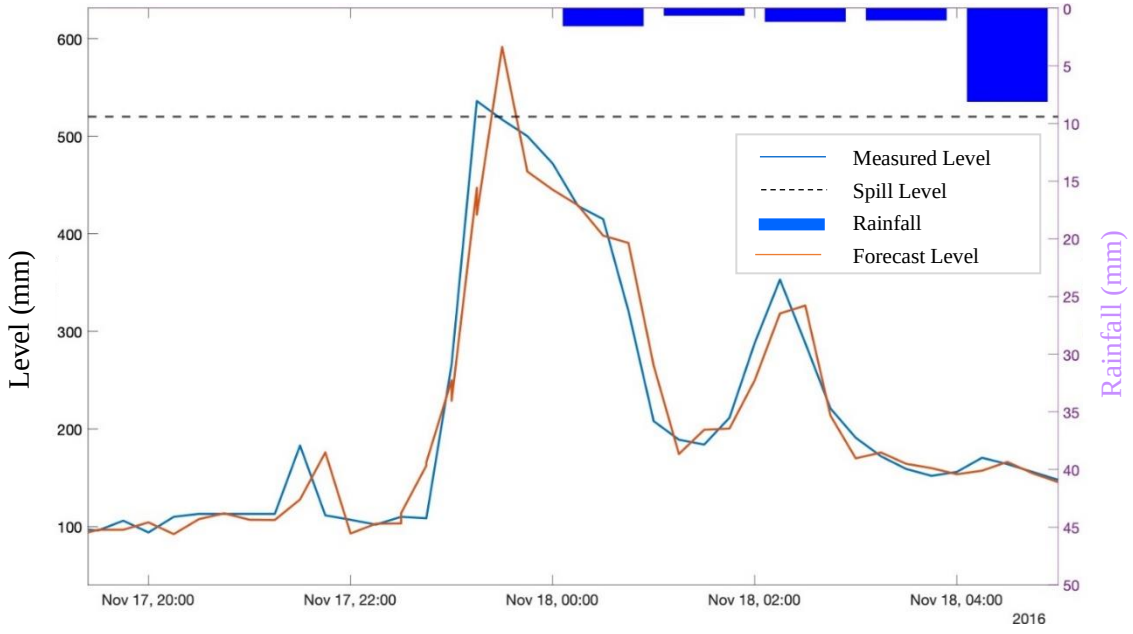


Figure 5-21 Measured CSO level vs CEANN Forecast level 2 hours ahead during an overflow event.

minute data, many of the actual spills are only a few timesteps in duration (34% of spills are 3 time-steps or less in length). Therefore, if the timing of the CEANN forecast is even one timestep out the Jaccard index will be very low. An example is presented in Figure 5-21 - a spill occurs from 23:15 to 23:30 and is forecast by the CEANN model from 23:30 to 23:45. As there is no overlap between the predicted and forecast spills the Jaccard index is 0. However, this forecast would still be valuable to a water utility and this should be reflected by the Jaccard index measurement. Therefore, by including the tolerance period T2 these slightly offset predictions are accommodated.

An analysis into the effects of varying the value of tolerance *T2* is presented in Figure 5-22. Values were considered from 0 to 3 hours (12 timesteps) for each forecast horizon of the CEANN model. Changing the value of *T2* is demonstrated to affect the raw Jaccard index number but has little effect on the trend of the Jaccard index over the range of forecast horizons. Therefore, we are confident that adding this tolerance does not significantly affect the conclusions obtained from the results of the analysis.

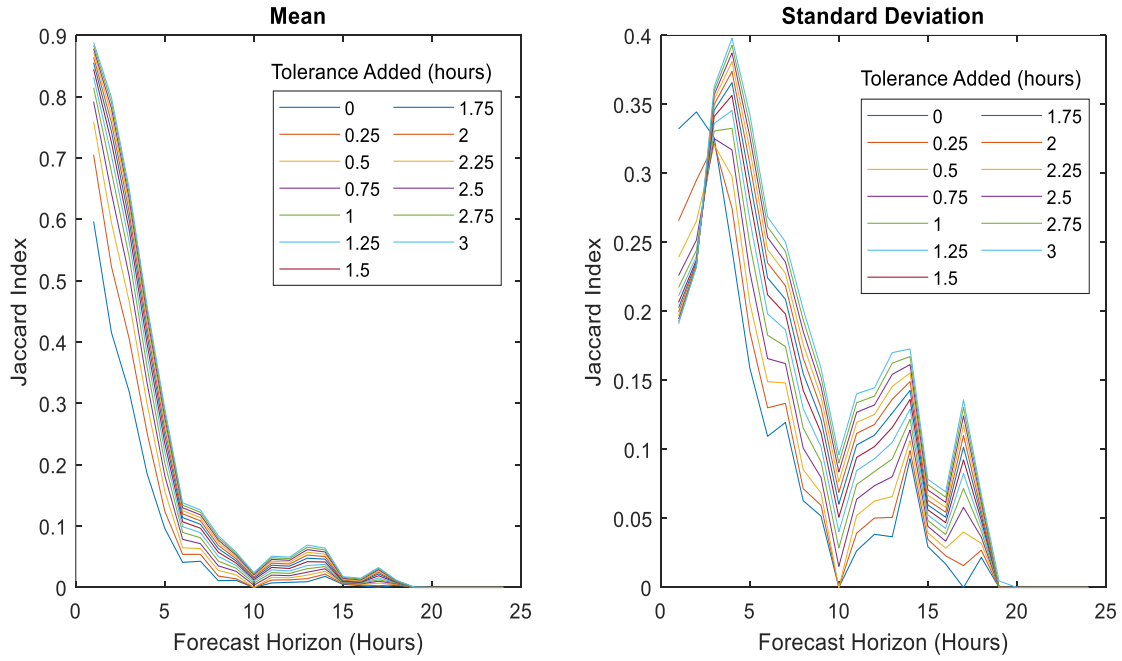


Figure 5-22 Effect of varying the tolerance time T2 on the Jaccard index mean and standard deviation.

The Jaccard index was computed for each identified spill event, and the mean and standard deviation over all the events were then calculated. Where a spill is not detected by the model the Jaccard index for that event is set equal to 0. The results for the two performance evaluations are reported in Table 5-7 for selected forecast horizons. As can be seen the ability of the CEANN model to detect the spills is very good at low forecast horizons (lead times of 2 hours and under). However, the performance deteriorates at higher horizons, and the number of true alerts produced when forecasting 2.5 hours ahead or more is very low. The same pattern is present in the Jaccard index results – the mean Jaccard index is high, up to one hour ahead, but decreases when predicting further into the future. The false alarm rate is generally very good for all the forecast horizons; the majority of forecast horizons produce no false alarms, only 5 horizons produce 1 or 2 false alarm over the 7.5 months of data the CEANN model was applied to.

Table 5-7 Results of the CEANN model overflow prediction evaluation for selected forecast horizons for T1 = 2 hours and T2 = 1 hour.

Forecast Horizon (hours)	Total Spills Starts Detected	Total Spill Starts Not Detected (i.e. False Negatives)	Spill Start Detection Rate	Total False Alarms	False Alarm Ratio	Jaccard Index	
						μ	σ
0.25	32	0	1	1	0.030	0.921	0.174
0.5	30	2	0.909	1	0.032	0.798	0.343
0.75	26	6	0.813	0	0	0.744	0.370
1	20	12	0.625	0	0	0.552	0.440
1.5	18	14	0.529	0	0	0.435	0.422
2	15	17	0.455	0	0	0.355	0.402
2.5	2	30	0.061	0	0	0.123	0.245
3	2	30	0.061	0	0	0.077	0.188
3.5	1	31	0.030	1	0.5	0.056	0.156
4	1	31	0.029	1	0.5	0.037	0.125
4.5	1	31	0.029	0	0	0.053	0.153
5	2	30	0.061	0	0	0.041	0.147
5.5	1	31	0.030	0	0	0.068	0.171
6	2	30	0.059	2	0.5	0.054	0.171

5.7.3 Alarm Level Threshold

The two performance analyses presented above demonstrate that the CEANN model can effectively forecast spill events and generate an alarm up one hour ahead. However, beyond this forecast horizon the spill prediction is poor. These spill alarms are still of use to a utility, but earlier detection is desirable to ensure that operators have enough time to adequately respond to an upcoming event.

An alarm threshold of 520 mm (the spill level of the CSO chamber) initially appeared reasonable. However, an examination of the level forecast data during the missed spill events indicates that the model generally predicts the increase in CSO level, but it does not reach the value of 520 mm required to generate an alarm. This is illustrated in Figure 5-23 which displays the measured CSO level vs the CEANN model forecast two hours ahead during three separate overflow events.

The use of the CEANN model, in comparison with the other three ANN models evaluated above, significantly improves the level predictions during wet weather, however, the model still has a tendency to underpredict the CSO level during the heaviest rainfall events (i.e. those that are likely to cause a spill) at high forecast horizons. This is compounded by the fact that, when the CSO level rises above the spill height, it generally is only by a small amount – the average CSO level during the 32 identified spills is 545 mm, in comparison to the spill height of 520

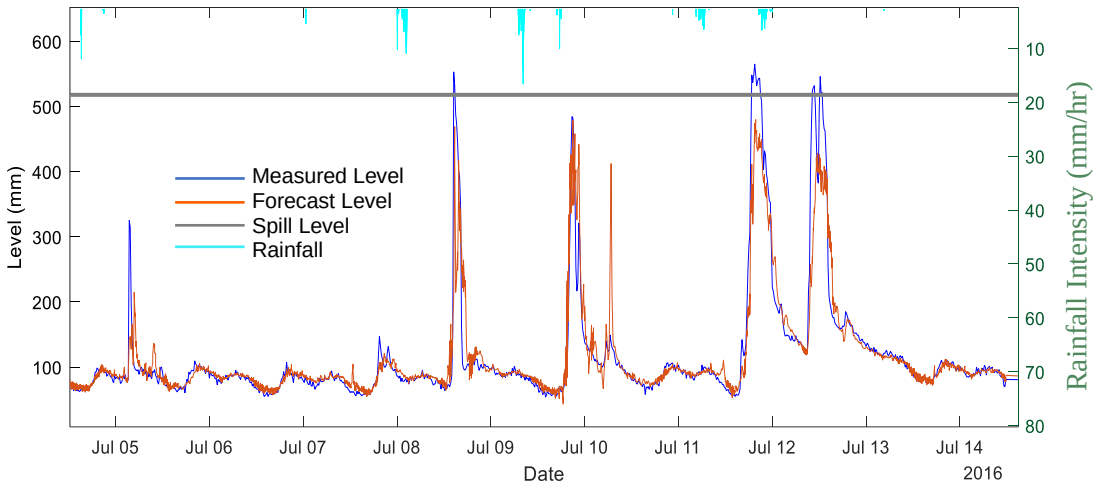


Figure 5-23 Measured CSO level vs CEANN forecast level 2 hours ahead during three overflow events.

mm. Thus, even a small degree of under-forecasting results in the system missing a large proportion of spill events.

To overcome this problem, an investigation into decreasing the forecast threshold required to generate alarm has been conducted, rather than using the actual spill height of 520mm. With the knowledge that the CEANN generally under-forecasts heavy rainfall events and does not tend to over-forecast, it is hypothesised that this should not generate a high number of false positives. Alarm threshold values were considered up to 520 mm (i.e. the actual spill level of the chamber). 60% of the dataset (14/2/18 to 03/7/18, containing 21 spills) was used for calibration of the alarm threshold and the remaining 40% of data (from 03/7/18 to 1/10/18, containing 11 spills) was used for testing the selected values on unseen data. The results presented below are for unseen data only.

Ideally, improving the accuracy if the CEANN model would be the optimal approach to improving the performance of the spill prediction methodology. However, adjusting the threshold is a simple way to generate better results.

Generally, an ROC curve would be used to describe the true positive and false positive rates that occur when varying the threshold value and provide a means for selecting the optimum cut off values. However, the construction of an ROC curve requires the calculation of number of false negatives. Alarms here are only generated when a spill is first detected, rather than at every timestep containing a spill. It is therefore preferable that the true negatives are not calculated here as, firstly, the duration of the spill is still of interest, and, secondly, tolerance periods have been included. Therefore, to consider all time periods not containing a spill start would be too strict a measure.

Therefore, instead of applying an ROC curve, the Jaccard index has been used to select the optimum alarm thresholds, which requires only the number of true positives, false positives and false negatives. When expressed in these terms the Jaccard index is calculated as

$$Jaccard\ index = \frac{True\ Positives}{True\ Positives + False\ Positives + False\ Negatives}$$

The process for identifying the optimal alarm thresholds is as follows. The alarm threshold values which produce the maximum Jaccard index for each forecast

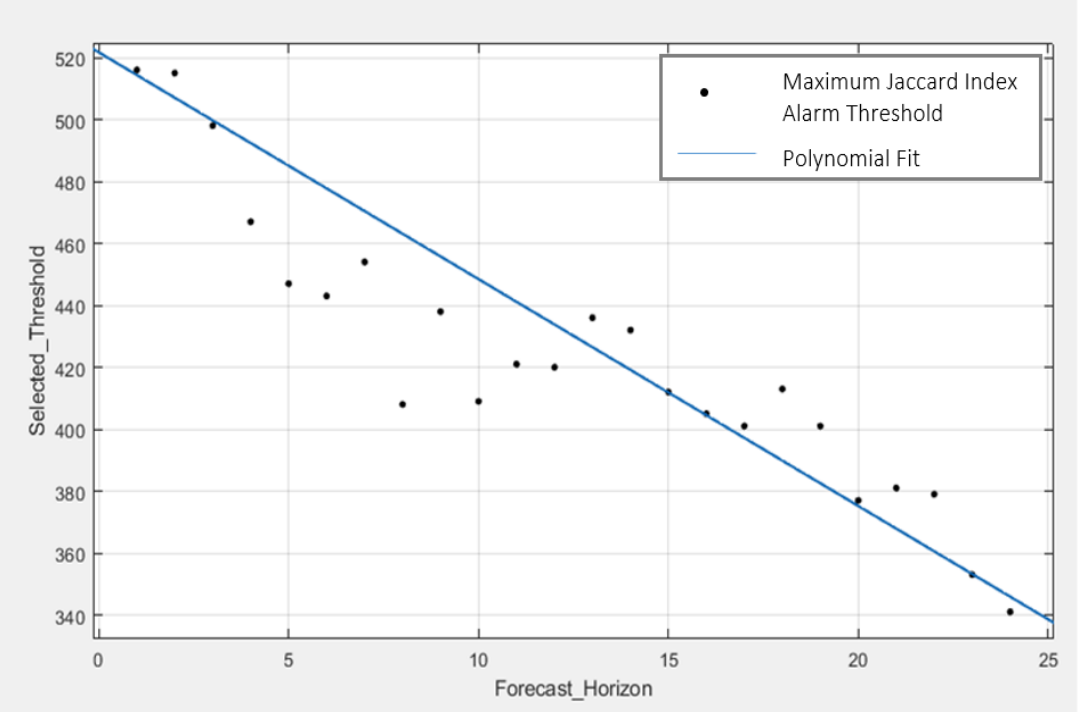


Figure 5-24 Linear model mapping forecast horizon to alarm threshold. Each point represents the identified threshold for each forecast horizon.

horizon are first identified. As discussed earlier, it is desired that the number of false alarms generated by the system are minimised, whilst still endeavouring to maintain a good true positive rate. However, as the Jaccard index is not a cost-based approach it cannot account for the different costs of true and false positives and the reduction of false positives cannot be emphasised when selecting the alarm threshold values. To overcome this, when identifying the maximum Jaccard index thresholds, all values with a total number of false alarms greater than 4 were excluded. This is a simple but robust method of controlling for the number of false alarms and was found experimentally to be effective. When utilised by a utility, the maximum allowable number of false alarms can be modified by the end user, based on their requirements.

A linear regression curve is then fitted to map the forecast horizon to these identified optimal alarm thresholds, using a least squares fit with increased weighting for the smaller forecast horizons, presented in Figure 5-24 This weighting was included as the smaller horizons are most affected by false alarms if the threshold is too low. This linear model is then used to select the final alarm

Table 5-8 Results of the CEANN model overflow prediction evaluation using a variable alarm threshold on unseen test data.

Forecast Horizon (Hours)	Selected Alarm Threshold	Total Spills Starts Detected	Spill Start Detection Rate	Total False Alarms	False Alarm vs True Alarm Ratio	Jaccard Index	
						μ	σ
0.25	520	13	1	0	0	0.951	0.058
0.5	507	13	1	0	0	0.843	0.287
0.75	500	12	0.923	0	0	0.836	0.275
1	492	11	0.846	0	0	0.811	0.363
1.5	478	11	0.846	0	0	0.731	0.400
2	463	8	0.615	0	0	0.689	0.410
2.5	448	7	0.500	1	0.125	0.452	0.387
3	434	5	0.357	2	0.286	0.507	0.361
3.5	419	6	0.400	2	0.250	0.445	0.389
4	405	8	0.571	2	0.200	0.361	0.414
4.5	390	7	0.500	1	0.125	0.265	0.376
5	375	11	0.733	2	0.154	0.224	0.347
5.5	361	10	0.667	7	0.412	0.411	0.404
6	346	13	0.867	6	0.316	0.296	0.376

threshold values. This model is simple enough that it should be able to generalise well to unseen test data for this CSO site.

The selected alarm thresholds for each forecast horizon using this method are presented in Table 5-8. As expected, the alarm thresholds decrease as the forecast horizon increases. The table also presents the results of the spill detection performance evaluations using the selected thresholds when applied to the unseen test dataset. The spill predictions are significantly improved - the proportion of true positives is considerably higher for all forecast horizons than when using the actual spill level as a threshold, as reported in Table 5-7.

The number of false positives has also increased slightly for forecast horizons between 2.25 and 5 hours, however they are still acceptably low (2 or under). The number of false spills generated for lead times of 5.5 to 6 hours is high, so it is recommended that either these forecasts horizons are not used to generate spill

alarms, or they are classed as low confidence alerts. On the calibration dataset the threshold was selected to limit the number of alarms to 4, but we note that when applied here to the validation dataset more than 4 alarms were produced for forecast horizons 5.5 and 6.

Overall, therefore, this spill alert methodology utilising variable alarm thresholds appears to perform well, although it has currently only been applied to one CSO site. Further tests will have to be conducted on additional CSO sites to analyse its performance further. The variable alarm thresholds will be specific for each CSO site the CEANN model is applied to, and therefore the procedure described above for identifying the optimal alarm values will have to applied performed for each site. However, the methodology is generic and can therefore be run automatically without requiring any additional input by a human operator. Therefore, applying this variable threshold method rather than using the single alarm value method described above is advantageous.

5.7.4 Section Summary

The main focus of the case study presented in this section is an evaluation of the ability of the CEANN model to effectively forecast upcoming overflow events and to generate alerts for upcoming spills. It has been shown that by using a variable alarm threshold for different forecast horizons the model can generate alerts in a timely manner and forecast the duration and timing of the spills well, whilst producing a low number of false alarms up to 5 hours ahead.

5.8 Summary and Conclusions

In this chapter the capabilities of the novel CSO level prediction CEANN model presented in Chapter 3 have been tested and evaluated on real case study CSO sites in the wastewater network. This investigation involved applying the CEANN model to historical level data from 4 CSO chambers of varying sizes and catchment characteristics in a simulated online fashion. The models were used to forecast CSO levels and predict upcoming overflows up to 6 hours ahead.

In order to analyse the CEANN models performance it was compared to 3 other ANN models - an ANN model developed manually through trial and error (i.e. the

Trial-and-Error ANN, approach adopted in existing literature), a single EANN model designed for all weather conditions (i.e. the All-weather EANN), and a joint wet and dry EANN model combined using a discontinuous transition (i.e. Combined Wet/Dry EANN model).

After the introduction and an overview of the objectives of the case studies, Section 5.3 described the case study sites data and Section 5.4. evaluated the performance of the data-pre-processing methodology, particularly it's ability to remove benching, Section 5.5 then described an initial study of an ANN model which aimed to determine several design parameters, namely (i) an investigation into the selection of the OS grid squares used for rainfall data inputs, and (ii) an analysis into the effect of including rainfall nowcast data as a model input. Next, in Section 5.6 the CEANN model and the three comparison models were applied to the four CSO case study sites and their performances evaluated. Finally, Section 5.7 described an evaluation of the ability of the CEANN model to forecast spill events and produce alarms for the upcoming events.

The results obtained from these analyses lead to the following conclusions:

1. The use of forecast rainfall data measurably improves the accuracy of the CSO level predictions using an ANN model for all forecast horizons analysed. This is noteworthy as nowcast data is observed to become less accurate at higher forecast horizons. However, the case studies demonstrate that the improvements gained are most significant at high forecast lead times even though this is when the nowcast data is least accurate.
2. The CEANN model produces more accurate CSO level predictions than the other three ANN models. This is true for all forecast horizons (up to 6 hours) clearly demonstrating the superiority of the CEANN methodology, especially over the Trial-and-Error ANN often used in the existing literature.
3. The CEANN and the Combined Wet/Dry EANN model accurately forecast the dry weather and wet weather level, predicting the timing and magnitude of the spill, thus providing information that is of clear use to a wastewater utility in near real time. The same cannot be said for the other two models (Trial-and-Error ANN and All-weather EANN) that continually

underestimate the CSO level during periods of heavy rainfall and hence cannot predict overflows.

4. The CEANN methodology is generic and thus requires minimal human effort to design and can automatically be applied to different catchments and forecast horizons.
5. The CEANN model can be used to effectively produce alarms for upcoming spill events up to 5 hours ahead.

The above demonstrates that the CEANN model presented in chapter 3 can be used beneficially by wastewater utilities to model CSO levels in the wastewater network in near real-time and provide alerts for upcoming spills - enabling better decision making and proactive management of overflow events. As the model is designed to utilise only CSO level and rainfall data, which wastewater utilities routinely collect in real time, it can be easily applied across their network with no extra costs involved.

In the following chapter the capabilities of the Blockage Detection System, which makes use of the CSO prediction model evaluated here, is tested and evaluated on real blockage events recorded in the wastewater network.

Chapter 6: Case Studies for Blockage Detection

6.1 Introduction

This chapter describes data analyses carried out on case studies blockage events. The main aim of these analyses is to test and evaluate the performance of the novel methodology designed for the detection of blockages in wastewater networks, as described in Chapter 4. The detection methodology was applied to historical level data from 10 CSO sites in the United Utilities sewer system containing blockage events.

The chapter is organised as follows. First a description of the case study sites and blockage events is given in Section 6.2. An evaluation of the performance of the detection system when applied to the case studies is then presented. The most relevant factors used the assess the performance of the detection system are the true positive, false positive and false negative rate, and the detection time of the blockage events. Next an in-depth illustration of the results obtained by the detection system when applied to individual CSO sites is presented in Section 6.3. These sites were selected to illustrate the performance of the system when applied to different types of events. The case sites studied consist of a sudden blockage event, a gradual blockage event, a CSO site experiencing sever benching, and a site experiencing multiple blockages. Finally, a summary of the chapter and the main conclusions are given in Section 6.4.

6.2 Case Studies Overview

This section presents an overall evaluation of the EDS performance when applied to real blockage events which occurred in the United Utilities sewer network.

6.2.1 Case Study Data

Sixteen blockage events, from 10 different CSO sites were designated for testing of the detection methodology. These blockages form the set of real-life events used to evaluate the performance of the detection system on unseen data.

All the blockage events were identified via a manual visual inspection of the historic level data to identify anomalous behaviour with the characteristics of a blockage. This was found to be the most reliable method of identifying blockage events. The identified blockages were then compared with United Utilities blockage removal report data from the surrounding catchment to determine if the blockage was removed by the wastewater utility. This report contains information on all the blockages removed from the United Utilities sewer network, recording the date of removal, the blockage location, the blockage cause, the sewer response (e.g. internal flooding), the method used to confirm the presence of the blockage (e.g. CCTV), the method used to remove the blockage and the overall cost of the procedure. However, it is important to stress that the location recorded for a blockage event is generally the location of the customer complaint which alerted the utility to its presence, rather than the actual location of the removed obstruction. Therefore, the actual location of the blockage in the sewer network is not known precisely.

In order to identify the blockage timings (i.e. the start and end times) a careful visual inspection was of the CSO level data was conducted. The end time of blockages are generally clearly discernible in the data, as the obstruction are largely either removed by the wastewater utility or dislodged by heavy rainfall, resulting in a very rapid decrease in CSO levels to normal behaviour. Blockage start times are more ambiguous – especially if the blockage began to form during rainfall, when the CSO level in the chamber is fluctuating. Therefore, the given times should be regarded more as an estimate. The starts times of gradually forming blockages, which cause CSO levels to increase very slowly over a long duration are particularly ambiguous. Therefore, when identifying the start time of gradual blockages only the day is identified, in comparison to sudden blockages which identify the exact time. This should be considered when analysing the detection system results.

The sixteen blockage events designating for testing of the detention system were selected to be representative of different characteristics (e.g. a mixture of gradual and sudden blockages, various blockage causes, durations, and sewer response). The CSO chamber sites were also selected to be diverse (e.g. different locations, catchment size, CSO chamber size and spill frequency). A summary table describing the characteristics of the case study sites and the blockage events is presented in Table 6-1. The blockages described here all resulted in an increase in CSO level. Blockages which cause a level decrease in the CSO chamber do occur too (e.g. blockages in sewers upstream of the CSO); however, they are significantly rarer and therefore have not been considered here.

Only one of the ten CSO chambers is located in a rural catchment. An attempt was made to include more rural blockages; however, few rural events were identified in the available historical CSO level datasets. Urban areas are known to be at a significantly higher risk of experiencing blockages and many blockage risk analysis methods and studies take into account land use classification in their methodologies (e.g. Bailey, 2016; Wallerstein & Arthur, 2012). Therefore, that the majority of the blockage events analysed here are urban is unsurprising. Note that the increased impervious cover in urban areas decreases the amount of rainwater that can naturally infiltrate into the soil and increases the volume and rate of stormwater runoff. This leads to more frequent and severe flooding in general, and so increases the likelihood of flooding during a blockage event. It is therefore important that utilities are alerted to the presence of urban blockages in a timely manner.

Of the 16 blockages events, 3 are gradually forming, and the remaining 13 formed suddenly. Gradual events appeared less commonly in the analysed historical datasets, however, they manifest differently in the sewer chamber and provoke a different level response therefore it is important to demonstrate that the detection system can identify both types of events.

The mean CSO dry weather level before and after the blockage event are reported to give an indication of the severity of the event. Blockages which cause the CSO level to rise above the spill height of the chamber, such as all three blockages which occurred at CSO site 2, are likely to have caused the CSO to

Table 6-1 Summary table of case study sites' characteristics.

CSO Site ID	Location	Sewer Network Type	Blockage Event Number	Blockage ID	Blockage Type	Blockage Duration	Dry Weather Spill During Blockage	Average dry weather CSO level before spill (mm)	Average dry weather CSO level after spill (mm)	Spill Level (mm)	Blockage Removed by Utility
1	Wirral	Urban	1	1	Sudden	6 days 5 hours	Yes	70	1036	1000	Yes
2	Cumbria	Urban	1	2	Sudden	32 days 4.5 hours	Yes	55	164	147	Yes
			2	3	Sudden	66 days 12 hours	Yes		160		No
			3	4	Sudden	11 days 19.25hours	Yes		159		Yes
3	Cheshire	Urban	1	5	Sudden	27 days 2 hours	Yes	125	195	210	No
4	Wirral	Urban	1	6	Sudden	34 days 7 hours	No	13	51	150	No
			2	7	Gradual	335 days 20 hours	No		31		No
5	Liverpool	Urban	1	8	Sudden	2 days 20 hours	Yes	58	467	900	No
			2	9	Sudden	5 days 6.5 hours	No		499		No
6	Greater Manchester	Urban	1	10	Sudden	4 days 11.25 hours	No	31	198	565	No
			2	11	Sudden	31 days 14.5 hours	No		368		No
7	Cumbria	Urban	1	12	Gradual	229 days	No	21	70	150	No
8	Greater Manchester	Urban	1	13	Sudden	8 days 15.75 hours	No	17	151	680	No
			2	14	Sudden	12 days 8 hours	No		435		No
9	Carlisle	Urban	1	15	Sudden	43 days 8 hours	Yes	88	869	550	Yes
10	Cumbria	Rural	1	16	Gradual	51 days 1 hour	Yes	27	104	390	No

overflow for the majority of the event duration, indicating an extremely serious event. The occurrence of a dry weather spill is also an important indication of event severity. Dry weather spills are particularly detrimental to the surrounding environment and potentially harmful to human health as they contain undiluted sewage. They can also incur serious penalties from the regulator. Eight of the 16 blockage events caused a dry weather spill during the time period the blockage was present in the sewer system.

Only 4 of the blockages identified in the data correspond to blockages recorded as removed by the wastewater utility. Regarding the remaining 12 blockage events, this implies that either the utility was not aware of their presence or that they were not significant enough to warrant any remedial action. Discussion with industry personal suggested that for the majority of these blockages it was likely that they were unaware of their presence, as any obstructions which cause dry weather spills should be immediately removed.

The EDS was applied to the level and rainfall datasets for each CSO case study site. Information regarding the source and frequency of the datasets can be found in Chapter 5. The dataset utilised for calibration, training and testing of the system

Table 6-2 Case study sites datasets.

CSO Site ID	EANN Model Calibration	Blockage Detection System Training	Blockage Detection System Testing
1	1/5/2016 – 1/8/2016	2/8/2016 - 1/3/2017	2/3/2017 - 18/7/2017
2	1/1/2016 – 1/3/2016	2/3/2016 - 14/7/2016	5/7/2016 - 8/2/2017
3	1/3/2016 – 1/5/2016	1/9/2016 - 12/2/2017	2/5/2016 - 1/9/2016
4	18/1/2016 – 18/3/2016	19/3/2016 - 18/8/2016	19/8/2018 - 6/11/2017
5	21/2/2016 – 21/4/2016	19/10/2016 - 12/5/2017	22/4/2016 - 18/10/2016
6	26/2/2017 – 25/5/2017	23/10/2017 - 11/5/2018	26/5/2017 - 22/10/2017
7	1/7/2016 – 26/10/2016	8/3/2018 - 25/5/2019	27/10/2016 - 7/3/2018
8	1/1/2017 – 10/5/2017	11/5/2017 - 25/9/2018	26/9/2018 - 25/5/2019
9	21/1/2016 – 21/3/2016	22/3/2016 - 22/8/2016	22/8/2016 - 1/12/2016
10	1/5/2016 – 27/7/2016	28/7/2016 - 18/2/2017	18/2/2017 - 18/10/2017

are presented in Table 6-2. In general, 20% of the level and rainfall datasets were used to train the EANN model required for the EANN discrepancy analysis subsystem, 40% of data was used for calibration of the event detection system and the remaining 40% (which included any blockage events) was used for testing of the system. However, the precise percentage varied for each case study site, due to differences in the duration and timings of the blockage events and due to other characteristics of the individual datasets.

During testing, the system was applied in a simulated real-time fashion to unseen data, i.e. as it would be used by a utility. The overall test data from all the 10 case study sites was 2240 days in total, of which 715 days occurred during a blockage event. The EANN models were required to have a minimum training data length of 2 months without blockages. A preliminary analysis showed that the precise length of data used to train the EANN model utilised in the EANN discrepancy-based analysis module did not measurably affect the performance of the overall detection system.

A single EANN model used to forecast the CSO levels was constructed for each site, as described in Chapter 3. A single EANN model was utilised instead of the superior bi-model CEANN described in Chapter 3, as the EANN model is less computationally expensive to construct. The CEANN model has been demonstrated to substantially improve level forecasts at high lead times, however when only predicting 15 minutes ahead the difference between the 2 models is not significant – as was demonstrated in Chapter 4. Therefore, there is not a measurable difference in the performance of the EDS using either the EANN or the CEANN model.

The parameters for the blockage detection system were selected during the methodology calibration, as described in Chapter 4. An alarm suppression period of 24 hours was applied. As described in Chapter 4 this suppression period suppresses any further alarms once an initial alarm has been raised in order to avoid raising unnecessary alarms for the same blockage event. An analysis into the effect of varying the value of this period is presented in Section 6.2.3.

6.2.2 Blockage Detection Results

The results of the blockage detection methodology when applied to the case study data are presented in Table 6-3. For each case study site the results are given in terms of the detection time of any blockage events (i.e. the time elapsed between the identified start of the blockage event and the generation of the first corresponding alarm), the total number of false alarms generated for that site and the rate of false alarms for that site (calculated as the number of false alarms

Table 6-3 Overall results of the blockage detection system.

CSO Site ID	Blockage Event Number	Blockage ID	Blockage Start Time (Ground truth)	Blockage First Detected by System	Time to Detection (Hours)	False Alarm Rate	Total False Alarms
1	1	1	11/07/17 21:00	11/07/17 22:45	1.75	0	0
2	1	2	21/07/16 11:30	21/07/16 14:15	2.75	5.97×10^{-4}	4
	2	3	8/11/16 16:15	15/11/16 12:00	163.75		
	3	4	28/01/17 23:45	29/01/17 21:45	10		
3	1	5	20/06/16 16:00	20/06/16 21:45	5.75	0	0
4	1	6	21/08/16 21:30	22/08/16 07:15	5	0	0
	2	7	20/10/16	15/11/16 21:00	645		
5	1	8	26/05/16 22:45	27/05/16 07:00	8.25	0	0
	2	9	13/08/16 06:45	13/08/16 22:15	5.5		
6	1	10	19/07/17 19:15	19/07/17 23:15	4	6.18×10^{-4}	6
	2	11	15/18/17 12:25	15/18/17 14:45	2.5		
7	1	12	07/11/18	12/11/18 20:30	140.5	0	0
8	1	13	27/10/18 23:30	28/10/18 12:00	12.5	1.12×10^{-4}	3
	2	14	07/11/18 11:45	07/11/18 14:45	3		
9	1	15	19/10/16 14:00	19/10/16 16:15	2.25	2.05×10^{-4}	1
10	1	16	12/09/17	19/09/17 21:00	189	0	0

generated by the system divided by the total number of ‘non-event’ timesteps for that site).

Overall, 100% of the blockage events were detected by the system – indicating that the methodology is effective at detecting different types of blockage events. The detection time of the system is also generally very satisfactory. Regarding the sudden blockage events, 12 of the 13 blockages were detected in under 24 hours and 9 detected in 6 hours or less. This appears to be a good result considering the average duration of the sudden blockages events (i.e. the time between the blockage occurrence and the blockage removal, due either to removal by the utility or heavy sewer flow due to rainfall) is in the order of 22 hours. Only the second blockage event from case study site 2 (i.e. blockage ID 3) has a comparatively long detection time of 163.75 hours. However, this is due to the CSO level response to the blockage event and is explained in further detail in Section 6.3.4. All the sudden blockage events were detected by the system before any dry weather spills occurred – therefore if applied by a utility in near real time the detection system would have alerted the company to the presence of the blockage before any pollution incidents occurred and may have prevented any unconsented spills.

Regarding the gradual blockage events (Blockage ID 7, 12 and 16), these blockages took longer to detect – at 645, 140.5 and 189 hours, respectively. However, this is to be expected as the change in CSO level over time caused by gradual blockages is generally very small. Discussion with industry personnel indicated that any preventative action (e.g. jetting) to remove a gradually building blockage would only be undertaken once the sewer flow was demonstrably obstructed and could potentially cause flooding during rainfall. Therefore, it is acceptable that the detection system is slower at identifying these types of events. The results of the detection system when applied to the gradual blockage event 2 from cases study site 4 is presented in more detail in Section 6.3.2 and illustrates that the system detected the blockage before a significant increase in the occurred.

A total of 14 false alarms were generated over the 10 CSO sites. This corresponds to a false alarm rate (i.e. the number of false positive timesteps/the total number of timesteps without a blockage) of 0.0089%. The ratio between

false positives and true positives can be viewed as relatively high, at 14 to 16. This is due partly to the nature of the dataset: the blockages events are rare (the vast majority of data is of non-blockage periods), and therefore the classification accuracy required to successfully detect the majority of events without incurring an equivalent number of false alarms would need to be incredibly high. Consequently, to ensure that the majority of blockage events are detected the number of false alarms generated is relatively high compared to the number of true positives. However, as stated, this corresponds to an overall low error rate when considering the total amount of data the system was applied to.

This false alarm rate of 0.0089% may appear exceptionally low. However, note that it represents the total number of false alarms (i.e. timesteps where a blockage is wrongly predicted) divided by the total number of timesteps not containing a blockage event. As the system is applied to 15-minute data and run over many months there is a very large number of non-blockage timesteps (many thousands). Therefore the false alarm rate is extremely small. This may falsely make the system appear to be performing extremely well. Other metrics were considered; however they did not take into account the amount of data that the system is tested on - i.e. communicate that 2 false alarms generated over 10 months of data is a better result than 1 false alarm generated over 1 month of data. Thus the false alarm rate was used in conjunction with the other metrics presented here.

The false alarms occurred on 4 CSO sites only. An examination of the false alarms showed that they generally occurred during wet weather, when the level was higher than ‘normal’ for that rainfall classification. For some sites it is possible that the sewer was partially obstructed by material during these periods, resulting in an increase in the CSO level, however this cannot be confirmed. This is the case for the 4 false alarms which occurred on CSO site 2 and is presented in Section 6.3.4.

CSO site 7, which generated the highest number of false alarms, was anomalous in that five of the six false alarms occurred during a timestep classified as ‘dry weather’. However, as shown in Figure 6-1 the level at these times remained high due to a previous rainfall event. As described in Chapter 4, a post event period

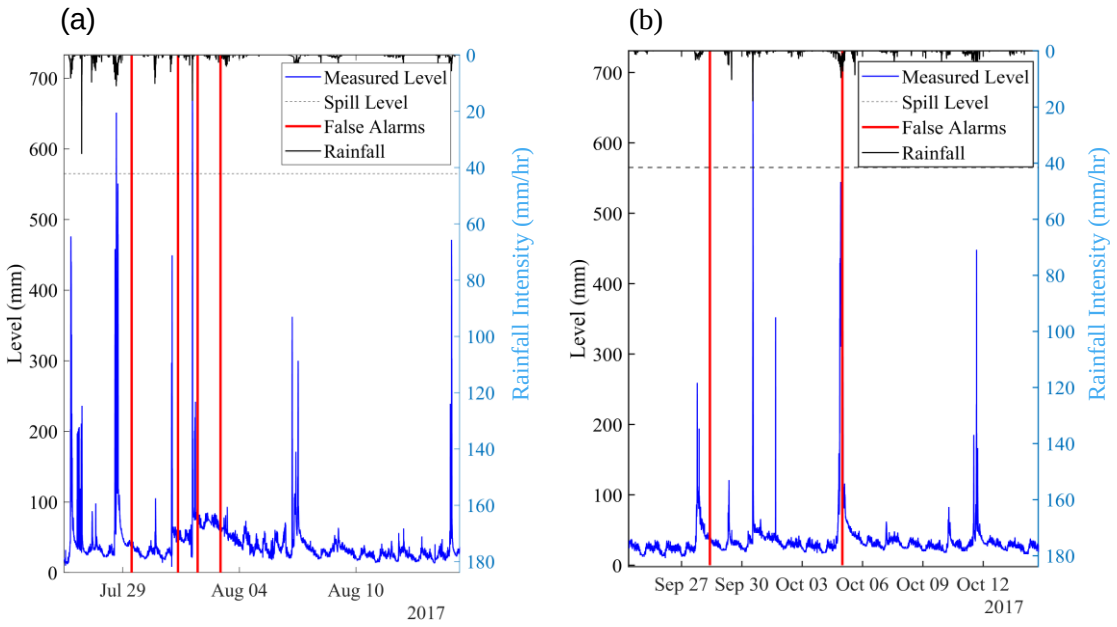


Figure 6-1 False alarms generated by the detection system for CSO site 7 during (a) July to August and (b) September to October.

was implemented to account for this period of increased level immediately following a rainfall event. The duration of this post event period is generic, selected based on an analysis of different CSO chambers and designed to accommodate all CSO sites. It appears, however, that this CSO chamber has a particularly long time of concentration and therefore the generic post event period is not long enough. When the post event period of the detection system was increased from 3 hours to 8 hours, only 1 false alarm occurred. Potentially future versions of the detection methodology could utilise an individualised post event duration, derived from the data for that site, rather than a generic value. However, there is benefit in have a generic parameter because the model requires less set-up. Additionally, this issue does not occur for the majority of sites, so it appears that this is anomalous.

Overall, therefore the detection system has been shown to have the capability to detect different types of blockage events reliably and in a timely manner, and with a low rate of false alarms.

6.2.3 Evaluation of the Alarm Suppression Period

The alarm suppression period is implemented by the detection system to avoid raising unnecessary alarms for the same blockage event. During this period the system suppresses any further alarms once an initial alarm has been raised, for a specified amount of time. The length of the suppression period is a user defined threshold, selected by operational personal. The results above utilised a 24-hour suppression period – this value was chosen assuming a single alarm per day was sufficient.

The selection of this period has the potential to influence the effectiveness of the detection system. For example, there is a possibility that the implementation of the period could result in missing events, e.g. if a false alarm is generated and a true event then occurs in the alarm inactivity period. Similarly, the suppression period could also result in an increase in the detection time. Therefore, the impact of applying this period, and the effect of varying the period length was investigated to aid the user in the selection of the most appropriate value. Different suppression period values were considered, from 0 minutes to 1 week and the effect on the true positive, the false positive rate and the detection time of the system were determined.

Overall, the analysis found that increasing the suppression time does not affect the true positive rate - all 16 blockage events were correctly detected by the EDS for all the various suppression periods. The detection time of the system was affected for only 2 of the 16 blockage events, blockage event 2 from case study site 8, and blockage 2 from case study site 6. Regarding site 8, blockage 1 ended on 3/11/2018 6:45 and blockage 2 began on 07/11/2018 11:45. Therefore as only 101 hours separate the events, when implementing an alarm suppression period greater than 101 hours the system cannot identify second blockage in a timely manner. Regarding CSO 6, a false alarm occurred 132 hours before the event, and so a suppression time greater than 132 hours increased the detection time of the event.

The outcome of varying the suppression time on the false positive rate is reported in Figure 6-2. The number of false alarms generated by the system decreased significantly when increasing the suppression period, from 1386 false alarms

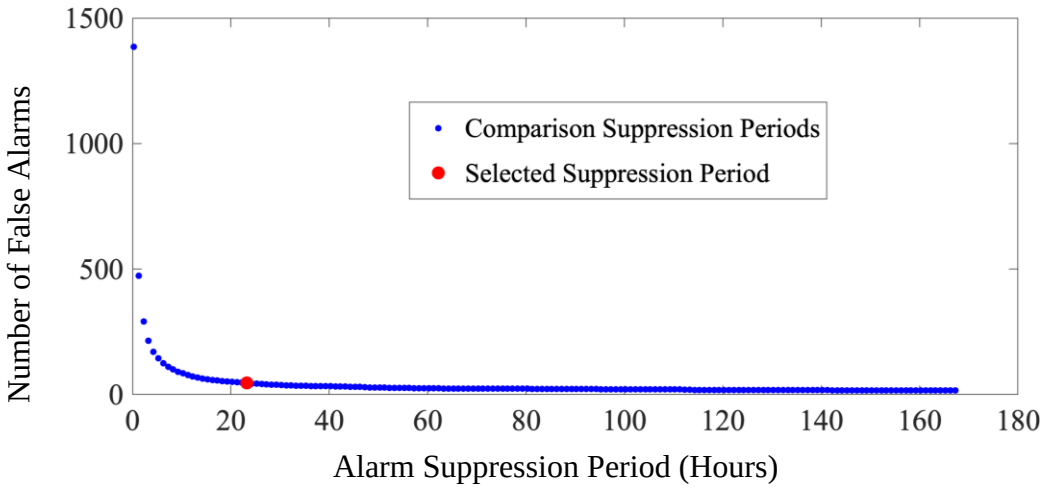


Figure 6-2 False positive rate for various alarm suppression period values.

when applying no suppression (a false positive rate of 0.99%), to 14 false alarms when applying the selected suppression period of 24 hours (a false positive rate 0.0089%), and to only 5 false alarms when applying a suppression period of one week (a false positive rate 0.0032%). This substantial decrease occurs as, when not applying any suppression period, false alarms often appear in 'batches', with many false alarms caused by unusual level behaviour over a short period of time. Therefore, applying the alarm inactivity period of only a few hours removes a large number of these false alarms.

The suppression period has therefore been shown to be effective at removing false alarms and generally does not affect the detection rate or detection time of the methodology. The threshold can thus be selected by the user based on company requirements for the acceptable alarm frequency.

6.2.4 Evaluation of the Inference Engine Modifications

The inference module is designed to raise an alarm when the submodules generate evidence of a blockage event, as explained in Chapter 4. The detection system has been demonstrated to be sensitive to the presence of blockage events – correctly detecting all 16 of the blockages presented above. However, the system does not discriminate based on the severity of the detected blockage. A wastewater utility using the system may not want to be alerted to blockages which only cause a small increase in sewer level and are interested only in events causing serious overflows. It is better not to generate a large number of alarms which will not be responded to. Additionally, small blockages can cause abnormal

behaviour which are detected by the above subsystem but are not obvious to the operators. In these cases, if a user believes the system is generating (too many) false alarms, they may lose trust in the system.

Therefore, the inference engine of the detection system was modified to raise an alarm at a timestep only if the level in the CSO is greater than a specified threshold height. This value was defined as a certain percentage of the spill height of the chamber. All other aspects of the detection system remained the same. The threshold was defined as above as this enables the use of a single, intuitive criterion which works well across different CSO chamber shapes and sizes.

A study was conducted to determine the effect of varying the level required to raise an alarm and to select a good threshold. For each case study site values were analysed from 0% of the chamber spill level (i.e. no CSO level requirement) to 100% of the spill level. The total number of blockages correctly identified by the system for different values is displayed in Figure 6-3. Applying the threshold level requirement does result in blockages not being detected for higher percentage thresholds – for example requiring a level in the chamber to be equal to the spill height results in only 9 (out of 16) blockages detected.

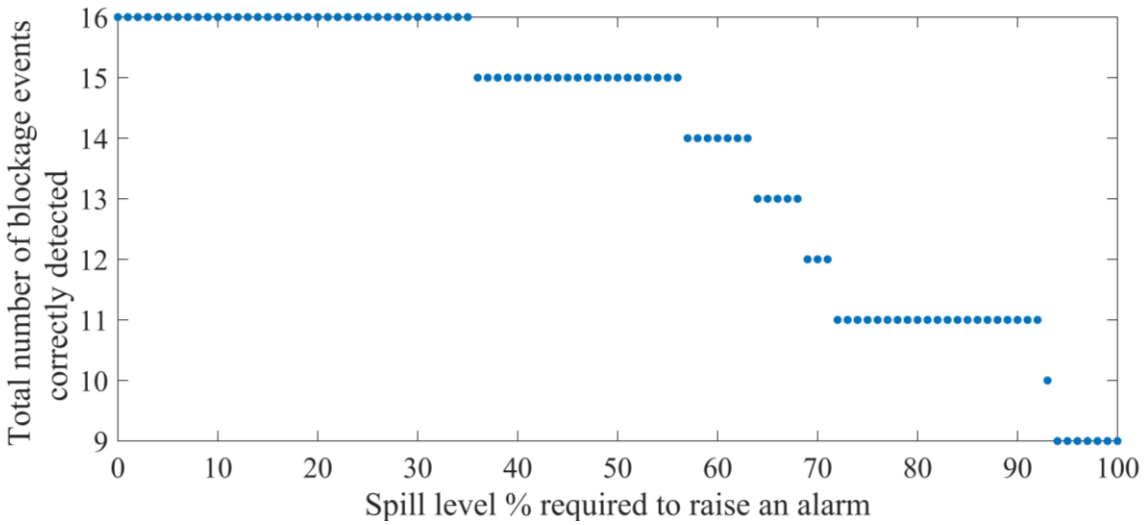


Figure 6-3 Total number of blockages detected by the system for different spill level percentage thresholds required to generate an alarm (out of a total of 16 blockages).

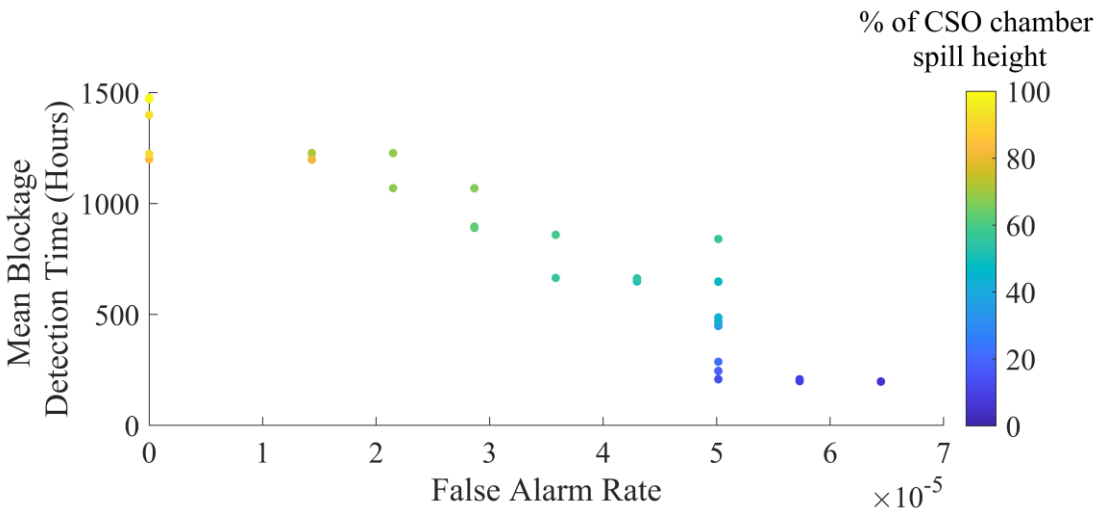


Figure 6-4 Mean blockage detection time vs false alarm rate for different spill height percentage thresholds required to generate a blockage alarm.

The overall false alarm rate vs the mean detection time of the system (i.e. the mean detection time of all 16 blockage events) for different thresholds is displayed in Figure 6-4. Increasing the threshold is demonstrated to increase the mean blockage detection time but reduce the number of false alerts. For example, requiring a level in the chamber to be equal to the spill height results increases the mean blockade detection time from 74.8 hours to 3621 hours.

An analysis of the detection system results revealed that generally it is the smaller blockages (i.e. blockages which result in a small increase in CSO level and so do not cause multiple spill events), which are not detected when increasing the threshold requirement. This is demonstrated in Table 6-4, which presents a comparison of the results obtained when requiring a CSO level of 30% and 50% of the chamber spill height before raising an alarm. Values in bold indicate those which have changed compared to the original results with no level requirement applied. As can be seen requiring a level of 50% spill height results in the two blockage events which occurred at CSO site 7 no longer being detected. Both events caused a comparatively small increase in level and did not result in a dry weather spill. The false alarm rate is also reduced, however, the detection time of certain events is increased.

Therefore, it appears that if a utility only wishes to raise alarms for significant blockages requiring a minimum level threshold is a viable option. This method decreases the sensitivity of the system resulting in the EDS not picking up smaller

events and with the benefit of reducing the false alarm rate. However, it also results in an increased detection time. The precise threshold required can be selected by the user based on their specific requirements.

Table 6-4 Comparison of the detection system results when requiring a level of 30% and 50% of the CSO chamber spill height to generate an alarm.

CSO site ID	Blockage Event	Blockage ID	No Minimum Level Requirement to Generate an Alarm		30% of Spill Height required			50% of Spill Height required
			Time to Detection (Hours)	Total False Alarms	Time to Detection (Hours)	Total False Alarms	Time to Detection (Hours)	Total False Alarms
1	1	1	1.75	0	**4**	0	**16.75**	0
2	1	2	2.75	4	2.75	4	2.75	**1**
	2	3	163.75		163.75		163.75	
	3	4	10		10		10	
3	1	5	5.75	0	5.75	0	9.75	0
4	1	6	5	0	**10**	0	**10**	0
	2	7	645		**3584**		**3584**	
5	1	8	8.25	0	8.25	0	**31**	0
	2	9	5.5		5.5		**10.25**	
6	1	10	4	6	4	0	**54**	0
	2	11	2.5		2.5		2.5	
7	1	12	140.5	0	140.5	0	**Not Detected**	0
8	1	13	12.5	0	9.25	0	**Not Detected**	0
	2	14	3	3	13	2	**255**	1
9	1	15	2.25	1	2.25	0	2.25	0
10	1	16	189	0	**201**	0	**226.75**	0

6.3 Detailed Case Studies

In this section a more detailed evaluation of the performance of the detection methodology on four CSO sites is presented. These events were selected to illustrate the capabilities of the detection system on different types of blockage events, and to demonstrate how the system could be utilised by a utility in real time. The detection system results presented here utilise an alarms suppression

period of 24 hours and include no level requirement to raise an alarm, unless expressly stated.

6.3.1 Case Study 1 - Sudden Blockage Event

This first case study describes the application of the blockage detection system to case study site 1, which experienced a sudden blockage event. This case study is presented to demonstrate how the system would benefit a wastewater utility if utilised in real time.

The CSO chamber in question is located in Merseyside, North West England in a predominantly residential catchment. A photograph of the CSO chamber, the CSO catchment and a diagram of the chamber and logger are presented in Figure

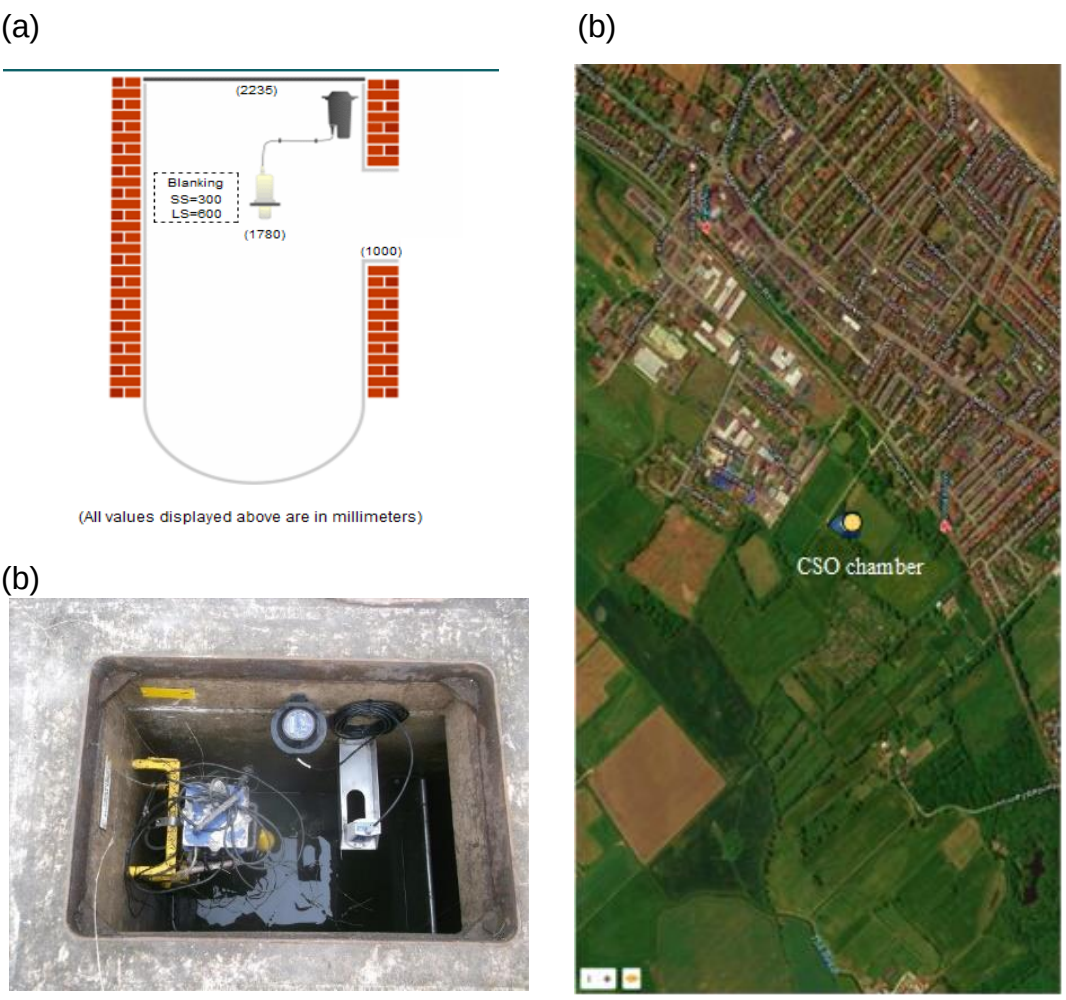


Figure 6-5 (a) Diagram of the CSO chamber, (b) photograph of the CSO level logger, and (c) photograph of the CSO chamber location for case study site 1.

6-5. An analysis of the level data from June 2016 to June 2018 reveals that the water level in the CSO chamber is above the spill height for 3.8 % of all timesteps, indicating that it does not spill overly frequently. There was some benching present in the data, which was removed during data pre-processing.

The historical CSO level data was analysed from 11/6/2016 – 01/08/17. One blockage event was identified in the 15-month dataset during a manual examination of the level data. The CSO level during this blockage is displayed in Figure 6-6. Based on a visual inspection the start time of the event was defined as 12/7/2017 and the end time as 18/07/2017. The CSO level during this period shows the classic characteristics of a sudden blockage event – a rapid increase as the blockage first obstructed the flow in the sewer, a plateau during which the diurnal patter is still visible, and lastly a sudden rapid decrease to normal levels as the blockage was cleared. The event lasted 6 days, overflowing continuously for the final 4 days. As this was an unconsented spill it had the potential to merit a penalty from the regulator. An analysis of the level data from nearby CSOs during this period showed that the blockage did not affect the level in any other CSO chambers.

The blockage event was cross referenced with the blockage removal report data from United Utilities for the area in proximity to the CSO chamber. A corresponding blockage removal was identified on 18/07/2017. Compared to the majority of the other blockages this obstruction was removed unusually

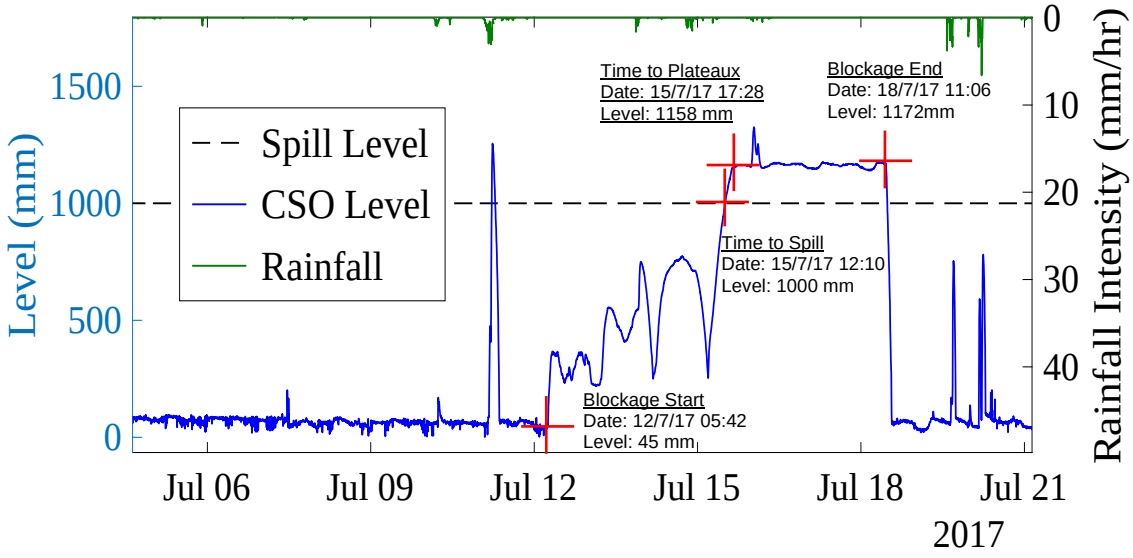


Figure 6-6 CSO level data during an identified sudden blockage event for case study site 1.

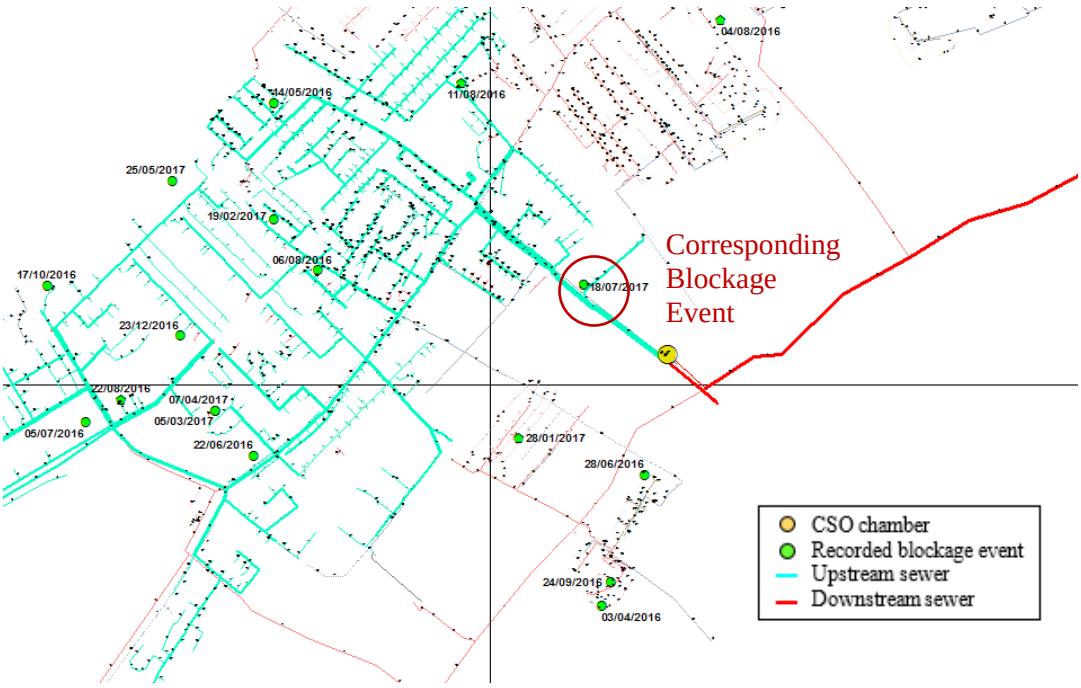


Figure 6-7 Blockages removed by United Utilities in proximity to the CSO chamber for case study site 1.

quickly – it can therefore be concluded that it had an obvious impact which required rapid removal. The reported location of the blockage from the removal report data can be seen in Figure 6-7. However, this is most likely the location of the customer complaint relating to the blockage event rather than the actual blockage location.

The information reported by United Utilities concerning the blockage removal is presented in Table 6-5. A soft blockage here refers to a build-up of soft materials flushed through the system, such as hair, grease, soap scum, or similar. This is in contrast to a hard blockage, which is caused by hard debris, such as a rocks or a dog's toy.

Many other blockages events were recorded as removed from the sewer system in the vicinity of the CSO chamber during the time period the CSO logger was installed in the chamber, as can be seen in Figure 6-7. However, a visual inspection of the level data during the dates identified for all the blockage removals showed that they did not correspond to any changes in the CSO level – indicating that the blockages did not affect the level in the CSO chamber. This

Table 6-5 Blockage removal information for case study site 1 recorded by United Utilities.

Cause	**Soft blockage**
Result	Surcharged system & Flooding
Response	Investigation & Jetting
Sewer pipe diameter	100 – 225 mm
Planned cost	£536
Actual cost	£745

is not uncommon if the blockage is at a sufficient distance from the CSO. This inspection process was repeated for all 10 case study sites analysed, with the same outcome. It can therefore be concluded that many blockages which occur in the sewer network unfortunately cannot be identified based on CSO level data.

This is an important limitation of the detection system. However, it is worth noting that the methodology was only applied here to data from CSO chambers as currently the majority of wastewater level monitors are installed in CSOs. In the future, as the price of monitors continues to decrease, monitors may become more widespread throughout the sewer system, increasing the potential range of the methodology.

The historical CSO level and rainfall data was collected from 11/6/2016 – 1/8/17. The EANN model was trained using 20% of the data, from 11/6/2016 to 1/9/2017. The event detection system was then calibrated using 50% of data, from 1/9/2017 to 1/4/2017, and the remaining 30% data was used for testing of the detection system, from 1/4/2017 to 1/8/2017, which contained the blockage event. The system was run in a simulated online fashion, i.e. as it would operate in real-time, for the test period dataset.

The results obtained by the detection system over the blockage period are displayed in Figure 6-8. The red vertical line indicates the timestep of the first alarm generated. The EDS detected the blockage event in 1 hour 45 minutes, 3

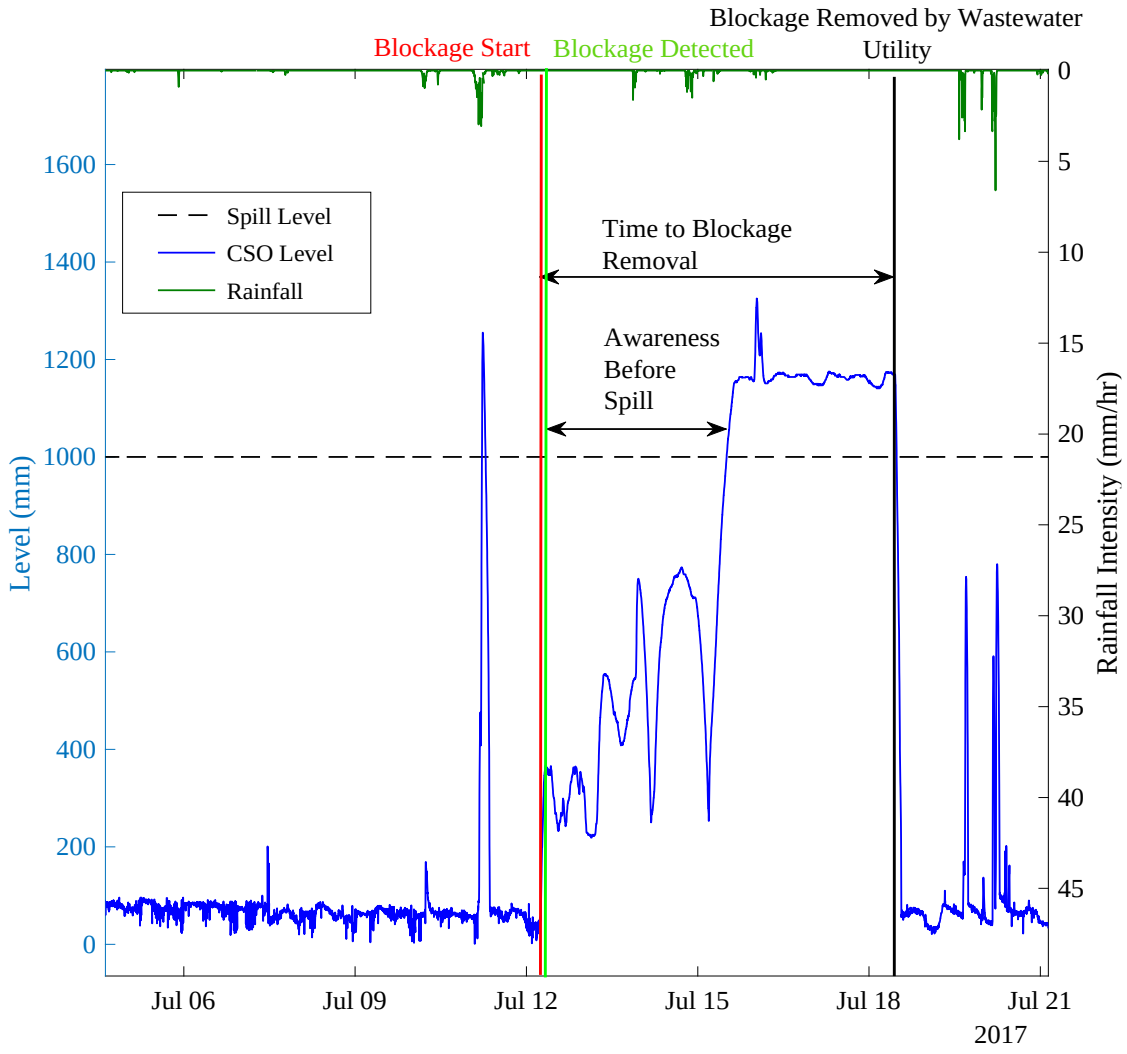


Figure 6-8 Results obtained by the detection system during an identified blockage event at case study site 1.

days before the CSO began to overflow and 6 days before it was removed by the utility. The blockage was detected whilst the increase in level was still small – showing that the system is sensitive to small changes in normal behaviour.

Over the 4 months of data used to test the system no false alarms were generated. Figure 6-8 displays a spill event which occurred on 11/07/2017 due heavy rainfall, the system recognised that this was normal behaviour and so did not generate an alarm.

Regarding the removal of the blockage by the wastewater utility; the site first began to spill during a rainfall event. Generally, alarms are raised automatically in the operations centre of a utility when a CSO overflows. However, the alarms

are commonly dismissed (either automatically or manually) when they occur during a rainfall event as they are assumed to be caused by the precipitation and the CSOs are therefore spilling within consent. This alarm dismissal is performed as during a heavy storm there may be hundreds of CSOs overflowing across the network – responding to each alarm to determine the spill cause would overwhelm the operators. It is likely therefore, that when this site began to spill utility personnel were not aware of the presence of the obstruction. It was only when the rainfall event ended and the CSO continued to spill that the utility was alerted to the problem. If the blockage detection system had been employed the utility, they would have been alerted to the blockage 3 days before it caused the unconsented overflow and may have been able to prevent the spill.

6.3.2 Case Study 2 – Gradual Blockage Event

The second case study focuses on CSO case study site 4, located in an urban catchment in The Wirral. This case study illustrates the performance of the detection system when applied to a gradually forming blockage event.

A CSO logger was installed in this CSO chamber on 17/03/2016, and level data was collected from 17/03/2016 to 07/11/2017. A visual inspection of the historic CSO level data identified two blockage events – a sudden blockage lasting 34 days and a gradual blockage lasting 335 days. An analysis of the blockage removal data from the surrounding catchment found no corresponding blockage removal reports for either blockage. However, the ends times of both blockage event correspond exactly with heavy rainfall events (shown in Figure 6-9),

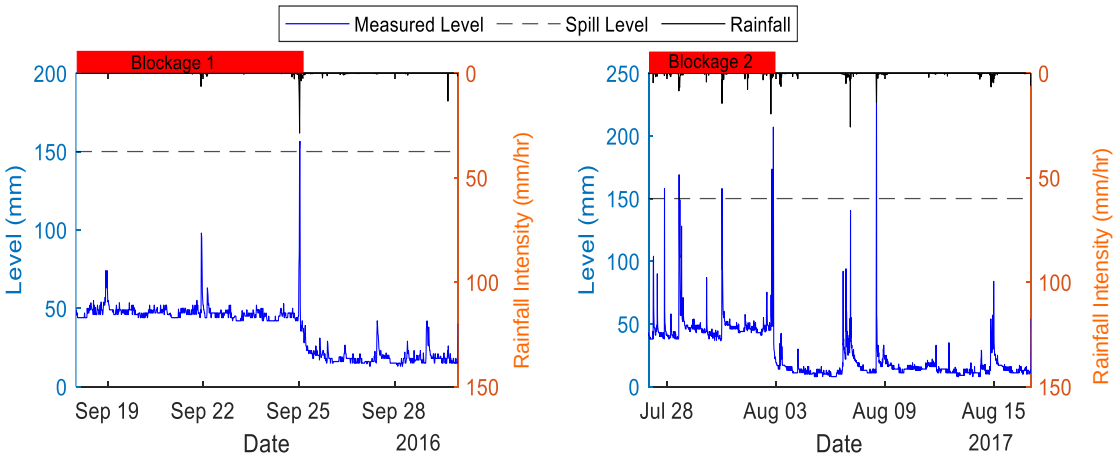


Figure 6-9 End times of the 2 identified blockage events at case study site 4.

indicating that the blockages were both dislodged by heavy flow due to rainfall. It is likely therefore that the utility was not aware of the presence of either obstruction.

Two months of data was used for training of the EANN model training (from 07/01/2017 to 07/01/2017). Of the remaining data, 25% of data, from 17/03/2016 to 17/08/2016 was used to calibrate the detection system and 75%, from 17/08/2016 to 06-Nov-2017 used to test the system. The comparatively small amount of data designated for testing was unavoidable as the two blockage events comprise such a large proportion of the dataset.

The applied to the historical test data set in an online fashion the EDS detected the gradual blockage event in 645 hours and the sudden blockage in 5 hours. The results obtained by the system during the gradual blockage event are presented in Figure 6-10. The green lines represent all the alarms generated by the system. As this is a gradual blockage the start time for this event should be regarded as an approximation only, it is more useful to look at the detection time of the system in relation to the change in CSO level data over the duration of the event. An alarm is first generated after only a small increase in level, the number of alarms generated then increases as the change in level becomes more significant. From April onwards an alarm is generated almost once every 24 hours

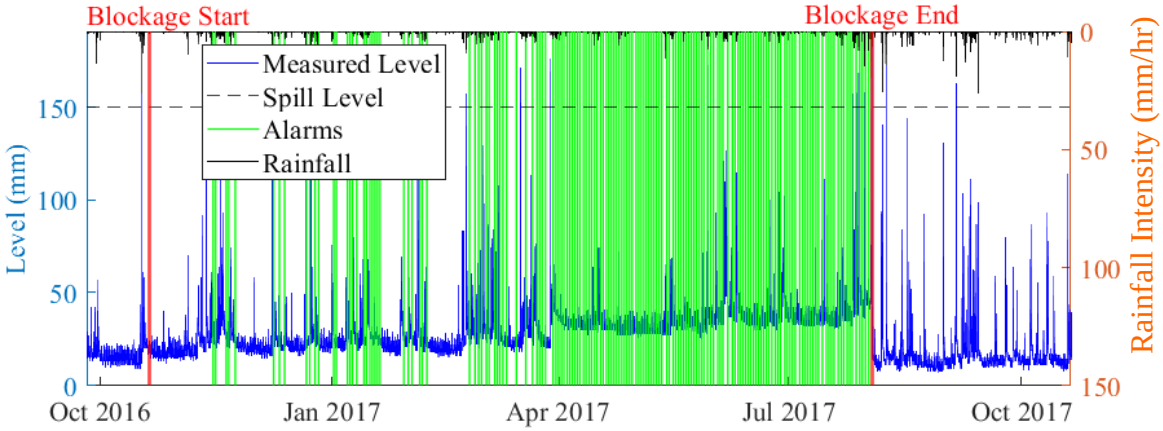


Figure 6-10 Results obtained from the blockage detection system for a gradual blockage event at case study site 4. The modified alarm refers to the first alarm generated by the modified system, which requires a level of 30% of the chamber spill height before an alarm is raised.

(note that the alarm suppression period is here set as 24 hours) indicating that every timestep is out of the bounds of normal behaviour. Employing the detection system would allow the utility to monitor the ongoing sewer behaviour before any spills occur and determine if removal is necessary.

The modified alarm shown in Figure 6-10 refers to the first alarm generated by the modified system (described in Section 6.2.4) which only generates an alarm if the level in the CSO is above a certain threshold. The threshold applied here is 30% of the spill height, although the result is the same if the level is set 50%. The detection time is delayed; however the blockage is still detected whilst the CSO level is still generally low.

6.3.3 Case study 3 – Occurrence of Benching

This third case study focuses on case study site 9, which experienced severe benching. As explained in detail in Chapter 3, benching refers to erroneous data points in CSO level data caused by the level monitor in the chamber mistakenly measuring the height of objects in the CSO chamber (such as the chamber wall) rather than the water level. Benching is identified and removed during the data-pre-processing.

The CSO chamber in this case study is located in an urban catchment in Carlisle. A single blockage event occurred during the available historical dataset which lasted for 43 days. The blockage was not removed by the wastewater utility. This case study is analysed to illustrate how the system performs when applied to a site with benching.

The level and rainfall data over the entire dataset is presented in Figure 6-11. The benching which occurs in the level data for this site is present only during the period the blockage is present in the sewer chamber – the recorded CSO level is normal preceding and following the blockage event. Initially it appeared that the logger here was malfunctioning, and no blockage had occurred. However, a closer inspection determined that this was a genuine increase in CSO level caused by a blockage as the diurnal sewer pattern can still be identified periodically in the level data during this period.

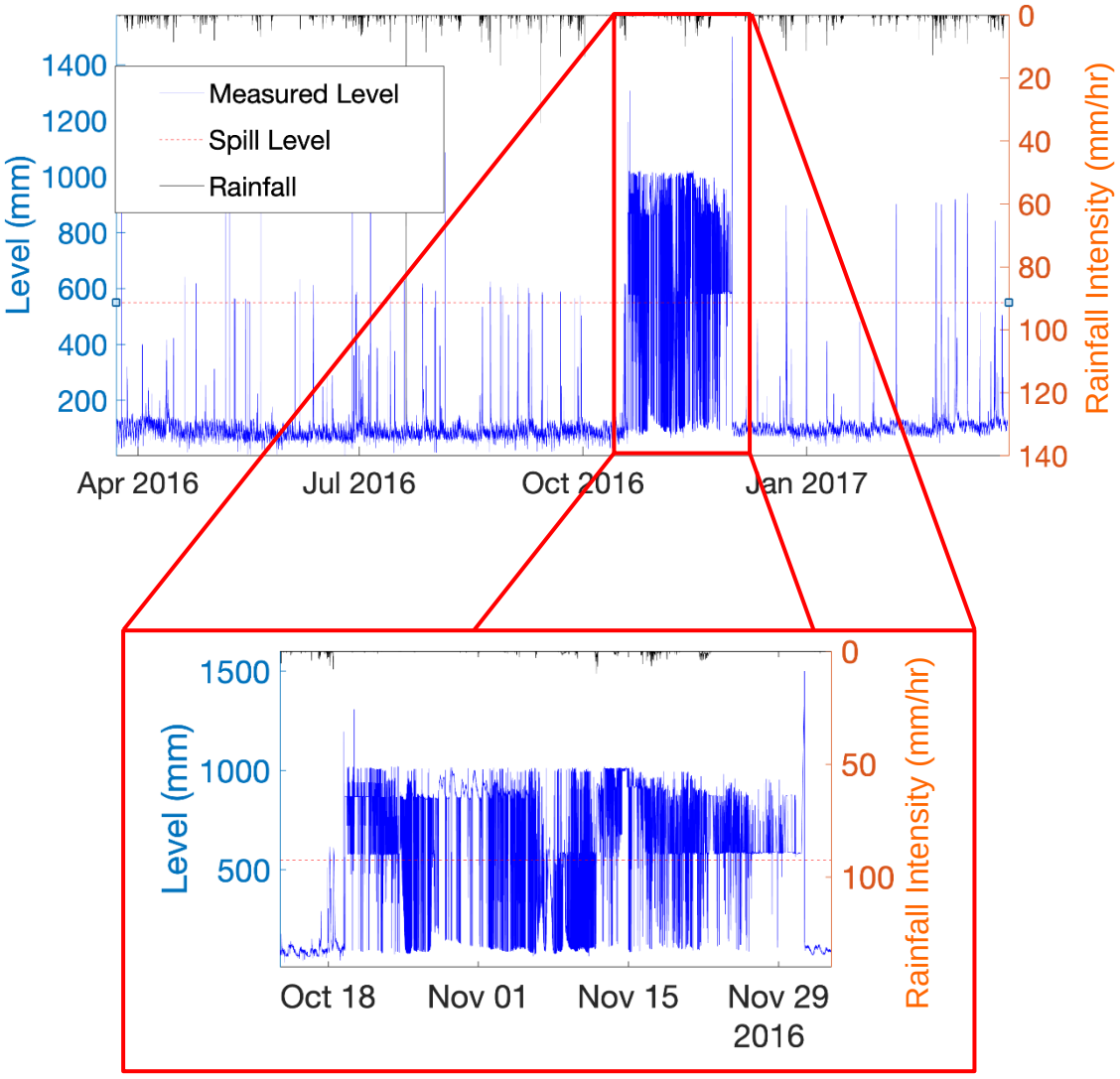


Figure 6-11 CSO level and rainfall data for case study site 9.

Benching removal was developed as part of the data processing methodology. However, identification and removal of benching is only applied to real time data if benching was initially identified during pre-processing of the calibration dataset. This is because benching generally occurs persistently throughout the dataset and so using the calibration information to direct the benching removal produces better results. As no benching occurred prior to the blockage event, this means the benching removal was not applied here to the real time level data. As benching of this sort is relatively rare, no other blockage events with this characteristic were available to use in the calibration of the methodology. As a result, the system was not designed to accommodate this sort of event and it was not considered during the selection of the system parameters. It was therefore unknown how the detection system would perform when applied to this site.

However, when applied to the test dataset the system identified the blockage event very quickly in 2 hours 15 minutes. It appears therefore that as the level is behaving abnormally the system was able to detect the blockage in a timely manner – even though it did not display the usual behaviour of a blockage event. One false alarm was generated over the 5 months of test data. Applying a level threshold requirement had no effect on the detection time of the system – even when a CSO level equal to the spill level was required to raise an alarm the detection time remained at 2.25 hours. However, an alarm threshold of 25% of the spill level or greater did result in no false alarms generated.

6.3.4 Case Study 4 – Multiple Blockage Events

The final case study focuses on case study site 2, which experiences multiple severe CSO spill events. The blockages which occurred here resulted in numerous dry weather spills but went undetected by the wastewater utility for many weeks. This case study aims to highlight the advantages that can be gained by applying this detection system to sites susceptible to frequent obstructions.

Seventeen months of historic CSO level and rainfall data was available for this CSO site, from 10/12/2015 to 02/05/2017. Data from 10/12/2015 to 10/02/2016 were used to train the EANN model, data from 10/02/2016 to 14/07/2016 was used for calibration of the detection system and the remaining data used for testing of the system. Three significant spill events were identified in the CSO level test dataset, shown in Figure 6-12. All three events caused multiple dry

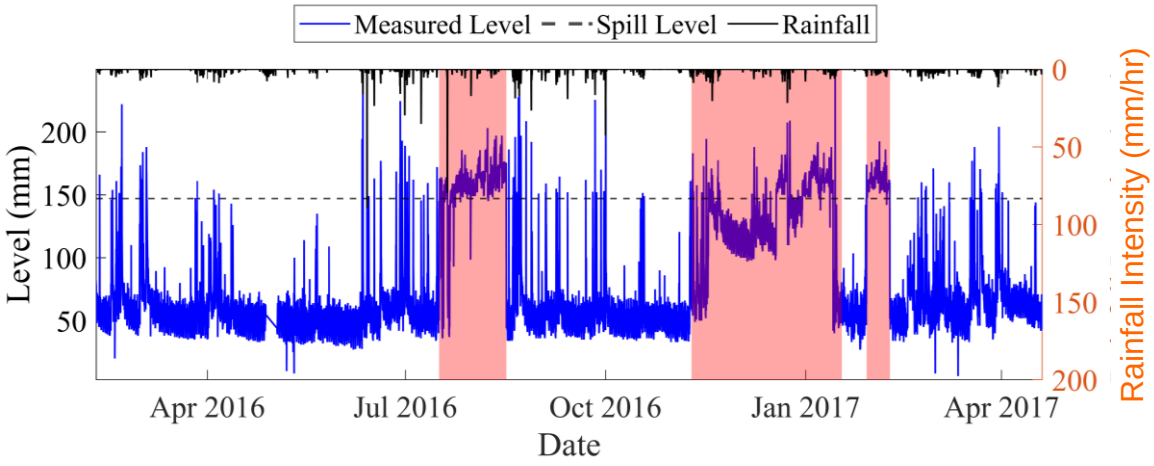


Figure 6-12 CSO level and rainfall dataset for case study site 2.

weather spills, indeed the CSO was overflowing for the majority of the blockages' duration.

According to United Utilities blockage removal data from the area surrounding the CSO chamber blockage 1 was removed by the utility 32 days after it started. The cause of the blockage was reported as a structural defect which resulted in a surcharged system. The presence of the blockage was confirmed using CCTV and an inspection probe. Blockage 3 was removed by the utility 11 days after it started, with the cause reported as a build-up of wipes or rags. Blockage 2, which lasted 66 days, corresponded with no removal data indicating it was not removed, and was eventually washed away by heavy sewer flow.

The results obtained by the detection system are reported in Figure 6-13. Blockage events 1 and 3 were detected quickly, in 2.75 and 10 hours respectively. Blockage 1 was detected most rapidly as it first manifested during dry weather, whilst blockage event 3 occurred during wet weather. Blockage event 2 has a comparatively long detection time at 163.75 hours. As can be seen

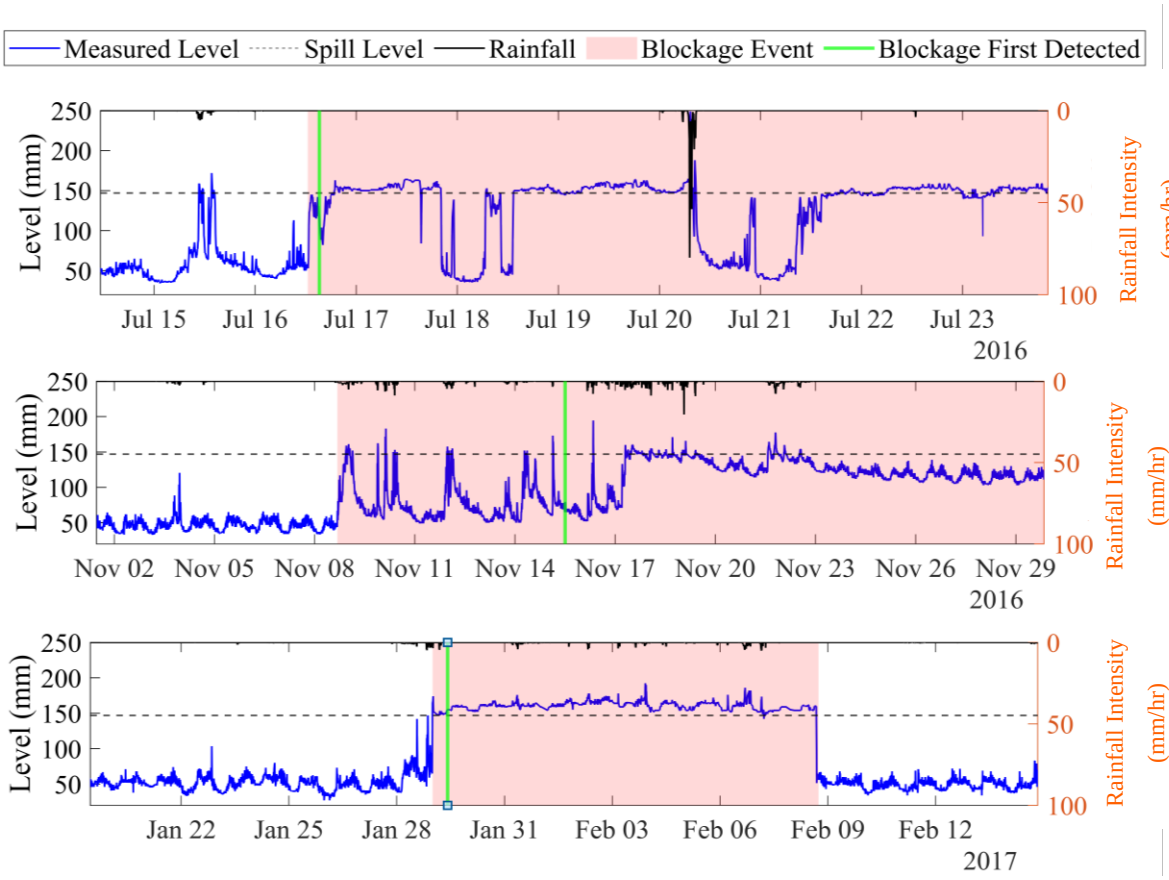


Figure 6-13 Results obtained by the detection system during 3 identified blockage events at case study site 2.

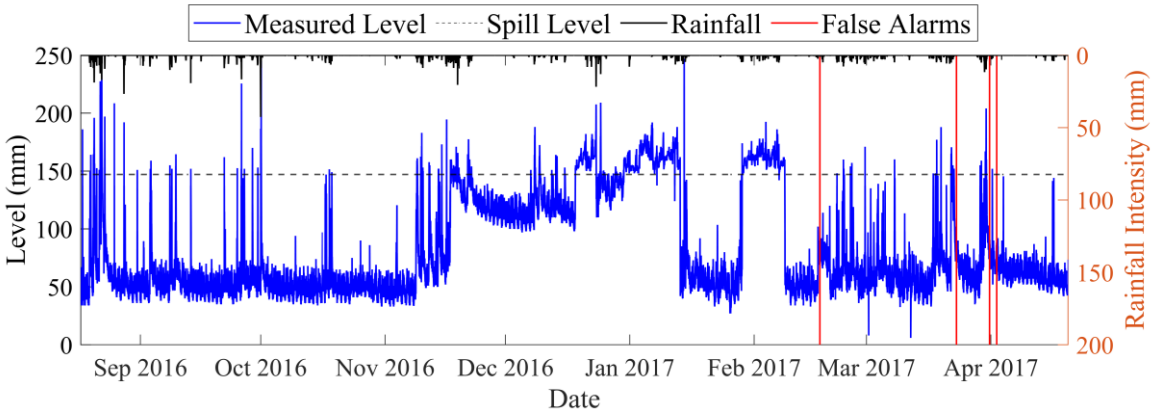


Figure 6-14 False alarms generated by the detection system for case study site 4.

in Figure 6-13, initially the blockage resulted in only a very small increase in level, from the 8th to the 17th of November. A much more significant increase then occurred from 17th November to the 14th January. It therefore took longer for the detection system subsystems to generate sufficient evidence to raise an alarm. However, it should be noted that the blockage was still detected a full 2 days before a long duration spill occurred.

Four false alarms were generated by the system over the test period dataset, displayed in Figure 6-14. All four alarms occurred during the period from March to May 2017. However, it is likely that these are due to 'mild' obstructions in the sewer network which cause a small change to the normal level behaviour. This is conjectured as the CSO level from March to May 2017 is visually more fluctuating than between September to November 2016, where no false alarms were generated, although the rainfall during the two periods is similar. Applying a threshold of 60% of the spill level height to raise an alarm generates no false alerts.

6.4 Section Summary

In this section the capabilities of the novel detection methodology presented in Chapter 4, which enables the detection of blockage events in the proximity of CSO chambers, has been tested on real blockage events situated in the United Utilities network. This was achieved by applying the detection system in a simulated real time fashion to historic level and rainfall data from 10 CSO chambers containing a total of 16 blockage events with varying characteristics.

After the introduction in Section 6.1, the overall results of the detection system when applied to the real blockage events was presented in Section 6.2. The system successfully detected all the blockage events and the false alarm rate was generally low – an overall rate of 0.0089% was obtained. Eight of the 11 sites the system was applied to produce no false alarms over the entire test dataset. The sudden blockage events were generally detected very quickly, before any dry spills occurred. The gradual events were demonstrated to have a longer detection time – however, this is expected as the increase in level due to gradual blockages is slow. Therefore, these obstructions are unlikely to cause flooding or pollution events soon after they manifest.

Section 6.3 then presented selected case study sites in more detail. These case studies were presented in order to illustrate the performance of the detection system on different types of blockage events. The case studies comprised of CSO site experiencing a sudden blockage event, a site experiencing a gradual blockage event, a CSO site containing significant benching, and a site experiencing frequent severe blockages and dry weather overflows. The methodology was demonstrated to have the ability to perform well for all 4 blockage types.

It should be noted that the methodology is only able to detect blockages which measurably disturb the level in a CSO chamber. An analysis of blockage removal data and CSO level data revealed that there are blockages in proximity to CSO chambers which are not captured in the level data. However, this limitation can only be overcome by installing additional monitors throughout the sewer system. The level data utilised here is CSO level data, as these are the most common monitors currently installed in the wastewater network. However, the methodology has the potential to be applied to data from any sewer level monitor installed in the wastewater network, without any need for substantial modifications. Therefore, as the number of loggers in the wastewater network increases, due for example to decreased costs, so will the detection range of the system.

The above demonstrates therefore that the detection system presented in Chapter 4 can be used to effectively detect blockage events in proximity to CSO chambers in a reliable and timely manner. Blockages in the sewer network often

go unnoticed for days or even months, as shown in these cases studies, and are only removed when flooding and other harmful consequences brings them to the attention of the utility or they are dislodged by increased flow due to heavy rainfall. The use of this system in real time has the potential to alert wastewater utilities to the presence of blockage events before they have a harmful effect, therefore decreasing potential damage to the surrounding environment, reducing clear-up costs and improving customer service.

The construction of the blockage detection system involved the selection of a large number of parameters, which required detailed consideration. As stated, it was desired that the methodology could be automatically applied to any site without requiring any individual preliminary analyses or parameter selection processes. It was therefore decided to utilise generic parameters only. The case studies have here demonstrated that generally this approach worked well when applied to real blockage data.

This generic approach can be applied to many other problems, in both urban drainage and water systems and in other fault detection applications. The selection process was accomplished here by applying the Youden index as well as a cost-based method, which allowed user-requirements to be integrated into the parameter selection. Although this required a lengthy initial selection process, once completed the results can potentially be used by wastewater personnel in the future to quickly and easily adjust the sensitivity of the system to produce the desired sensitivity. This approach could be easily transferred to similar data-driven fault detection systems in other fields.

In the following chapter a summary and discussion of the work carried out in this thesis is presented, in addition to the main conclusions gained and a discussion on the direction of further work on the topic.

Chapter 7: Conclusions

This thesis has presented the development of a methodology designed for the detection of blockages and other unusual events in the sewer network using real time CSO level data. Additionally, a prediction model has been developed, designed to forecast water level in a CSO. This chapter summarises and discuss these novel methodologies. It is organised as follows; firstly, a summary of the work presented in this thesis is presented. Next the key contributions of the thesis are described. The final section on future work then proposes a number of recommendations for future research and developments following from the contributions covered in this thesis.

7.1 Thesis Summary

The first chapter of the thesis presented an introduction to the work. A brief overview of historical wastewater management was presented, followed by a discussion of current blockage management practices, highlighting the motivation behind the work. The aims and objectives of the thesis were then defined, this includes the design of an ANN model capable of accurately forecasting levels in a CSO chamber and the development of an event detection system designed to detect blockages and other unusual event using only CSO level data and rainfall data. The methodologies must be generic and self-learning in nature, so that they can be applied to any CSO chamber, regardless of location, without requiring any input from a human operator.

Chapter 2 presented a review of the literature relevant to blockage detection in the wastewater sector. This included an overview of the current hardware and software techniques currently employed by wastewater utilities, and a discussion on fault detection methodologies and in particular the use of data driven models. The relevant gaps in the knowledge were identified.

The third chapter presented the first methodological chapter, describing the ANN model developed to forecast CSO levels in real-time and provide alerts for upcoming spill events is presented. First a data pre-processing methodology was detailed, designed to effectively process incoming CSO level and rainfall data.

Next the CEANN methodology was presented. This included techniques utilised to improve the model performance by overcoming data imbalance. In addition, three alternate ANN model were described, which were developed to evaluate the performance of the CEANN model.

Chapter 4 then described the development of the event detection system methodology. The system is designed to mimic the work of a trained, experience human technician in determining if a blockage event has occurred from the available CSO level data in an automated fashion. An overview of the EDS was first presented. This was followed by an in-depth description of the techniques designed to detect blockage events by identifying deviations in the CSO level data, namely a statistical analysis-based approach and an EANN discrepancy-based approach.

Chapter 5 evaluated and demonstrated the capabilities of the CEANN model on real CSO level data. The results were compared with the performance of the alternate ANN models and it was demonstrated that the CEANN model produces more accurate CSO level predictions for all forecast horizons. A number of analyses were then conducted to demonstrate the capability of the model to predict upcoming CSO spill events and provide alerts.

In Chapter 6 the capabilities of the event detection system were evaluated. The system was applied to data from several CSO sites containing a number of historical blockage events. The results obtained demonstrated that the system was able to successfully detect all the blockages reliably and in a timely manner.

7.2 Summary of Thesis Contributions

The work carried out in this thesis forms a useful contribution to the field of wastewater event detection. The key contributions are as follows:

- The development of a methodology for pre-processing of historic and real time CSO level data and rainfall data. This involved the development of a methodology to identify and remove benching from CSO level data.

- The development of an ANN methodology for the prediction of water level in a CSO chamber and the generation of alerts for upcoming spill events. Specifically, this involved:
  - The development of an ANN model utilising CSO level data, radar rainfall and forecast rainfall data. It was demonstrated that the inclusion of forecast rainfall data measurably improves the model predictions.
  - The development of an EANN model, whereby an evolutionary algorithm is employed to select the optimal model architecture and input structure.
  - The development of a bi-model CEANN, composed of two evolutionary artificial neural network models, optimised for wet and dry weather respectively, and combined using a non-linear weighted averaging approach in order to overcome bias arising from imbalanced data.
  - The novel application in this context of the structural similarity index (SSIM) as a better means of performance evaluation of the ANN prediction model
  - An investigation into the selection of an optimal threshold to generate spill alerts. This involved the application of the Jaccard index to analyse the quality of the predictions for different thresholds.

- The development of a statistical analysis-based methodology to detect blockage events in real time based on identifying deviations from normal CSO level behaviour. This involved developing:
  - A method for categorising incoming data based on rainfall duration and intensity.
  - A method for constructing statistical boundaries for each rainfall category within which observed level data should sit, assuming no event has occurred and the CSO level is exhibiting normal behaviour.
  - The selection of modified western electric control rules for each rainfall category to monitor, over consecutive timesteps, level data which falls outside the constructed boundaries indicating evidence of a blockage event. This involved the application of the cost index, a generalised version of the Youden index.

- The development of an EANN discrepancy-based methodology to detect blockage events in real time based on the deviations between short-term EANN model predictions and the observed CSO level data. Specifically, this involved the development of:
    - A method for constructing statistical boundaries for each rainfall category within which the EANN model prediction error should sit, assuming no event has occurred and the CSO level is exhibiting normal behaviour.
    - The selection of modified western electric control rules for each rainfall category to monitor, over consecutive timesteps, EANN error data which falls outside the constructed boundaries indicating evidence of a blockage event.

- The integration of the above methodologies into a single event detection system. This involved the development of an inference engine designed to combine the event detection evidence and determine if a blockage has occurred. The system was tested, validated and demonstrated on a number of real blockage events from UK case studies.

7.3 Thesis Conclusion

A novel detection system has been designed to detect blockages and other unusual events which occur in the wastewater network in the proximity of CSO chambers. The system is designed to operate in near real-time. In addition, a CEANN model has been designed to forecast CSO levels in real-time and provide alerts for upcoming spill events.

The key concept underlying the methodologies are their generic and self-learning nature – they can be applied to any CSO location without requiring any input from a human operator, providing a historical dataset of level and rainfall data is available for calibrating the system. Unlike physically based models, the methodologies do not require any physical understanding of the sewer system and are also much less expensive to build, calibrate and maintain. It is therefore anticipated that both the systems could be applied across a whole wastewater network, continually monitoring the level data for developing blockage events and predicting future overflows.

The blockage system has been demonstrated to detect all types of blockage events in an efficient and timely manner. Employing the system will allow utilities to respond promptly to developing blockage events, therefore reducing potential harm to the sewer network, and damage to the surrounding environment and properties due to flooding and overflows. The system will also improve the utilities operational performance and customer service.

7.4 Future Work Recommendations

7.4.1 Event Detection Methodology

The field of automatic real-time blockage detection is still new, and as a result there are many different opportunities for further development and expansion. Most significantly, the proposed detection system is early on in its development. Therefore, the most important future work is further testing and validation of the system on a large dataset, containing many types of different CSO sites, blockage events and operating conditions. It is important to determine if the current methodology is applicable to all CSOs, or if the methodology should be modified based on the site characteristics. The introduction of the Event Duration Monitoring project resulted in the installation of an increasing number of CSO level monitors in the sewer network. Therefore, it is likely that an ever-increasing amount of CSO level data will be available for testing of the system.

In addition to this, there are a number of different developments which could be investigated. These are described below:

- **Development of a blockage detection decision support tool.** A critical element for the deployment of the blockage detection system is the creation of a user-friendly decision support tool which can be integrated with the current technology utilised by wastewater utilities. This is vital to the successful application of the system in industry and should be created with the input of industry professionals.

- **Further development of the inference engine to determine alarm confidence.** There are many methods available to determine event confidence, the most commonly applied is Bayesian probability. Another

possibly would be to apply Platt scaling, which transforms the outputs of a classification model into a probability distribution over classes. This would allow personnel to rank incoming alarms.

- **Development of a methodology to determine the likely location of the detected blockage event.** An indication of the blockage location would be very beneficial for industry personnel when removing the identified obstruction. Localisation of leaks in wastewater distribution networks using pressure and flow data from sensors installed in the network is a popular area of research, and many of the developed technologies have the potential to be transferable to the wastewater network. These methodologies utilise machine learning and geostatistical techniques. Development of a blockage localisation methodology could be challenging since, as stated in this thesis, blockage removal data often does not contain the actual location of the removed obstruction and so acquiring a blockage location dataset to develop a methodology could be a difficult task. However, solutions are available – for example simulating blockage events manually by opening and closing valves in the sewer.

- **Development of a methodology identifying the likely root cause of a detected blockage event.** This would offer an improved interpretation of the system results and aid industry personal in making an informed decision when removing the obstruction. It has been established in this thesis that gradually forming blockages, caused by a build-up of siltation, and suddenly forming blockages, caused for example by snagging of hard objects, can be identified due to the different profiles apparent in the CSO level data. It is possible that blockage events causes could be further categorised, for example distinguishing blockages caused by sewer collapses, FOG, tree roots etc. This methodology would identify the blockage cause based on the characteristics of the level data deviations resulting from the obstruction. A preliminary analysis would be required to determine if different blockage types indeed create unique sewer level characteristics. Wastewater utility blockage removal records generally record the blockage type, therefore data would be easily available.

- **Application of the developed methodology to other data sources.** The core philosophy of the detection system is detecting deviations from normal behaviour. As the methodology is generic, transferring it to other data sources is possible. Applying the system to level data from sewer pipes (in comparison to CSO chambers) would be straightforward and require few modifications. The system could also be applied to data sources such as pumping station data.

- **Analysis of level data from multiple CSOs.** The presence of a blockage event in the sewer network can affect the level in multiple CSO chambers in close proximity to each other. Including data from multiple CSOs would increase the reliability and sensitivity of the system, providing additional evidence that a blockage has occurred if it is detected in level data from multiple loggers. Incorporating data from multiple CSOs would also be key in development of a blockage localisation methodology. This was not incorporated into the current methodology as insufficient data was available, however in the future this would be a promising direction.

- **Development of additional blockage detection subsystems**. As the developed methodology is modular, combining blockage evidence from different subsystems, including additional detection modules would be possible. For example, a methodology analysing the temporal patterns in the CSO data. The current detection methodologies analyse discrepancies from normal behaviour but do not analyse the actual changes in the level data during a blockage. It could be possible to identify a blockage event by training an ANN to learn the characteristics of the change in CSO level behaviour during a blockage event.

7.4.2 CSO Level Prediction Methodology

CSO level prediction using data driven models has longer history, however there is still scope for further developments. Potential directions include:

- **Increasing the prediction range of the model.** The current CEANN model has a forecast range of 6 hours, as this is range of the MET office

nowcast data. The MET office also generates short-range weather forecasts, from 1 to 3 days ahead, although at a lower temporal resolution. Therefore, a model forecasting a lead time of up to 3 days could be investigated. It would be useful to determine how far ahead the CEANN model is able to predict.

- **Inclusion of more data inputs to the CEANN model.** The model described in this thesis was intentionally developed utilising only CSO level data and rainfall data, so that it could be easily implemented by wastewater utilities without requiring them to invest in additional sensors. However, the inclusion of additional inputs could potentially improve the model results. Possible additional inputs include the soil moisture index, temperature, or humidity. Utilities often collect soil moisture data, although at a low temporal resolution, the data is available from the MET office at daily, weekly or monthly intervals, and temperature and humidity are widely available at high temporal resolutions.

- **Development of a recurrent ANN (RNN) model incorporating the concepts developed for the existing architecture.** The current methodology trains ANN models tailored to the specific CSO site and forecast horizon required. All the models are independent of one another. RNN models contain internal self-looped cells, allowing them to use previous outputs for making current predictions. Thus, the level forecast prediction for a lead time of 15 minutes could be utilised as an input for the 30 mins ahead model. RNNs are well suited to supervised learning problems where the dataset has a sequential nature such as time series forecasting. An RNN model could be developed incorporating the concepts developed for the existing model, such as the wet and dry models designed to overcome data-imbalance.

- **Development of a deep learning ANN (DNN) CSO level prediction model incorporating the concepts developed for the existing architecture.** DNNs utilise multiple layers to progressively extract higher level features from the raw input. Deep learning has made great strides in recent years, facilitated by the improvements in computer technology,

and some attempts have been made to apply the recent advances to hydrological problems. DNNs has been observed to perform better than shallower ones on some tasks. DNNs are prone to overfitting, because of the added layers of abstraction. This can be overcome by ensuring sufficient data is used to train the models. It is likely that a years' worth of data would be enough to yield good results. DNNs require more processing power than shallow models and the initial training of the models takes longer, which should be considered if applying the methodology to a large number of sites across the network. Additionally, the trained DNN models are generally larger (typically hundreds of megabytes in size) and require more processing power to run.

- **The development of a binary classification ANN model, solely for overflow prediction.** Training a model specifically for the task of overflow prediction has the potential for better performance than indirectly inferring the occurrence of spill events from forecast level data generated by the CSO level prediction model. This fits in with the philosophy of end-to-end training, whereby the entire process involves a single model without modularization, i.e. an end-end model will directly convert input data into an output prediction, bypassing the intermediate steps that usually occur in a traditional process.

In addition to all of the above, data-driven and machine learning technologies are continually and rapidly improving. Therefore, it is likely that new and more powerful techniques will be available for deployment in the field of event detection and wastewater management.

Prototype tools / software relating to the work carried out in this thesis will be made available upon request whilst respecting confidentiality of UU data.

References

Adamowski, J. S. (2008). Peak daily water demand forecast modeling using artificial neural networks. *J. Water Resour. Plann. Manage.*, *134*(2), 119–128. https://doi.org/10.1061/

Adams, B. J., Fraser, H. G., Howard, C. D. D., & Sami Hanafy, M. (1986). Meteorological Data Analysis for Drainage System Design. *Journal of Environmental Engineering*, *112*(5), 827–848. https://doi.org/10.1061/(asce)0733-9372(1986)112:5(827)

Al-Ani, R. R. A., & Al-Obaidi, B. H. K. (2019). Prediction of Sediment Accumulation Model for Trunk Sewer Using Multiple Linear Regression and Neural Network Techniques. *Civil Engineering Journal*, *5*(1), 82. https://doi.org/10.28991/cej-2019-03091227

Alfieri, L., Berenguer, M., Knechtl, V., Liechti, K., Sempere-Torres, D., & Zappa, M. (2015). Flash Flood Forecasting Based on Rainfall Thresholds. In *Handbook of Hydrometeorological Ensemble Forecasting*. Springer Berlin Heidelberg.

Alpert, P., Ben-Gai, T., Baharad, A., Benjamini, Y., Yekutieli, D., Colacino, M., Diodato, L., Ramis, C., Homar, V., Romero, R., Michaelides, S., & Manes, A. (2002). The paradoxical increase of Mediterranean extreme daily rainfall in spite of decrease in total values. *Geophysical Research Letters Banner*, *29*(11), 31-1-31–34.

Altobelli, M., Cipolla, S. S., & Maglionico, M. (2020). Combined application of real-time control and green technologies to urban drainage systems. *Water*, *12*(12), 1–15. https://doi.org/10.3390/w12123432

An, K., & Meng, J. (2010). Voting-averaged combination method for regressor ensemble. *Proceedings of the 6th International Conference on Advanced Intelligent Computing Theories and Applications*, 540–546.

Arthur, S., Ashley, R., Tait, S., & Nalluri, C. (1999). Sediment transport in sewers-a step towards the design of sewers to control sediment problems. *Proceedings of the Institution of Civil Engineers: Water, Maritime and Energy*, *136*(1), 9–19. https://doi.org/10.1680/iwtme.1999.31264

Arthur, S., & Burkhard, R. (2010). Prioritising sewerage maintenance using inferred sewer age: A case study for Edinburgh. *Water Science and Technology*, *61*(9), 2417–2424. https://doi.org/10.2166/wst.2010.176

Arthur, S., Crow, H., & Pedezert, L. (2008). Understanding blockage formation in combined sewer networks. *Proceedings of the Institution of Civil Engineers: Water Management*, *161*(4), 215–221. https://doi.org/10.1680/wama.2008.161.4.215

Ashley, R. M., Fraser, A., Burrows, R., & Blanksby, J. (2000). The management of sediment in combined sewers. *Urban Water*, *2*(4), 263–275. https://doi.org/10.1016/s1462-0758(01)00010-3

Azamathulla, H. M., Ab Ghani, A., & Fei, S. Y. (2012). ANFIS-based approach for predicting sediment transport in clean sewer. *Applied Soft Computing*, *12*(3), 1227–1230. https://doi.org/10.1016/j.asoc.2011.12.003

Bailey, J., Keedwell, E., Djordjevic, S., Kapelan, Z., Burton, C., & Harris, E. (2015). Predictive risk modelling of real-world wastewater network incidents. *Procedia Engineering*, *119*(1), 1288–1298. https://doi.org/10.1016/j.proeng.2015.08.949

Bailey, J. R. (2016). *Data-Driven Models of Blockage Likelihood in the Wastewater Network, PhD Thesis, Univeristy of Exeter*. https://core.ac.uk/download/pdf/83924203.pdf

Balbastre-Soldevila, R., García-Bartual, R., & Andrés-Doménech, I. (2019). A comparison of design storms for urban drainage system applications. *Water*, *11*(4). https://doi.org/10.3390/w11040757

Barzegar, R., & Asghari Moghaddam, A. (2016). Combining the advantages of neural networks using the concept of committee machine in the groundwater salinity prediction. *Modeling Earth Systems and Environment*, *2*(1). https://doi.org/10.1007/s40808-015-0072-8

BBC News. (2017). *"Monster" fatberg found blocking east London sewer - BBC News*. BBC. https://www.bbc.co.uk/news/uk-england-london-41238272

Benbassat, A., & Sipper, M. (2013). Evolving artificial neural networks with FINCH. *GECCO 2013 - Proceedings of the 2013 Genetic and Evolutionary Computation Conference Companion*, *87*(9), 1719–1720. https://doi.org/10.1145/2464576.2480780

Berenguer, M., Corral, C., Sánchez-Diezma, R., & Sempere-Torres, D. (2005). Hydrological validation of a radar-based nowcasting technique. *Journal of Hydrometeorology*, *6*(4), 532–549. https://doi.org/10.1175/JHM433.1

Berkhahn, S., Fuchs, L., & Neuweiler, I. (2019). An ensemble neural network model for real-time prediction of urban floods. *Journal of Hydrology*, *575*, 743–754. https://doi.org/10.1016/j.jhydrol.2019.05.066

Bischoff, R., & Guhl, T. (2010). The strategic research agenda for robotics in Europe. *IEEE Robotics and Automation Magazine*, *17*(1), 15–16. https://doi.org/10.1109/MRA.2010.935802

Bissell, A. F. (1978). An attempt to unify the theory of quality control procedures. *Journal of Applied Statistics*, *5*(2), 113–128. https://doi.org/10.1080/768370747

Boldt, M., Borg, A., Ickin, S., & Gustafsson, J. (2020). Anomaly detection of event sequences using multiple temporal resolutions and Markov chains. *Knowledge and Information Systems*, *62*(2), 669–686. https://doi.org/10.1007/s10115-019-01365-y

Borsányi, P., Benedetti, L., Dirckx, G., De Keyser, W., Muschalla, D., Solvi, A. M., Vandenberghe, V., Weyand, M., & Vanrolleghem, P. A. (2008). Modelling real-time control options on virtual sewer systems. *Journal of Environmental Engineering and Science*, *7*(4), 395–410. https://doi.org/10.1139/S08-004

Botturi, A., Ozbayram, E. G., Tondera, K., Gilbert, N. I., Rouault, P., Caradot, N., Gutierrez, O., Daneshgar, S., Frison, N., Akyol, Ç., Foglia, A., Eusebi, A. L., & Fatone, F. (2020). Combined sewer overflows: A critical review on best practice and innovative solutions to mitigate impacts on environment and human health. *Critical Reviews in Environmental Science and Technology*, 1–34. https://doi.org/10.1080/10643389.2020.1757957

Bougadis, J., Adamowski, K., & Diduch, R. (2005). Short term municipal water demand forecasting. *Hydrol. Processes*, *19*, 137–148. https://doi.org/10.1002/ hyp.5763

Bowler, N. E., Pierce, C. E., & Seed, A. W. (2006). STEPS: A probabilistic precipitation forecasting scheme which merges an extrapolation nowcast with downscaled NWP. *Quarterly Journal of the Royal Meteorological Society*, *132*(620), 2127–2155. https://doi.org/10.1256/qj.04.100

Breakey, D., & Meskell, C. (2013). Comparison of metrics for the evaluation of similarity in acoustic pressure signals. *Journal of Sound and Vibration*, *332*(15), 3605–3609. https://doi.org/10.1016/j.jsv.2013.02.033

Brokamp, C., Beck, A. F., Muglia, L., & Ryan, P. (2017). Combined sewer overflow events and childhood emergency department visits: A case-crossover study. *Science of the Total Environment*, *607–608*, 1180–1187. https://doi.org/10.1016/j.scitotenv.2017.07.104

BSI. (2000). Gravity drainage systems inside buildings. In *Bs En 12056-3:2000* (Vol. 3, Issue 1).

Butler, D., Digman, C., Makropoulos, C., & Davies, J. W. (2018). Urban Drainage. In *Urban Drainage* (4th ed.). Taylor and Francis Inc. https://doi.org/10.4324/9780203351673

Caloiero, T., Coscarelli, R., Ferrari, E., & Sirangelo, B. (2016). Trends in the Daily Precipitation Categories of Calabria (Southern Italy). *Procedia Engineering*, *162*, 32–38. https://doi.org/10.1016/j.proeng.2016.11.008

Carbone, M., Turco, M., Brunetti, G., & Piro, P. (2014). Minimum Inter-Event Time to Identify Independent Rainfall Events in Urban Catchment Scale. *Advanced Materials Research*, *1073–1076*, 1630–1633. https://doi.org/10.4028/www.scientific.net/amr.1073-1076.1630

Chaves, P., & Chang, F. J. (2008). Intelligent reservoir operation system based on evolving artificial neural networks. *Advances in Water Resources*, *31*(6), 926–936. https://doi.org/10.1016/j.advwatres.2008.03.002

Chaves, P., & Kojiri, T. (2007). Conceptual fuzzy neural network model for water quality simulation. *Hydrological Processes*, *21*(5), 634–646. https://doi.org/10.1002/hyp.6279

Chen, J., & Adams, B. J. (2005). Analysis of storage facilities for urban stormwater quantity control. *Advances in Water Resources*, *28*(4), 377–392. https://doi.org/10.1016/j.advwatres.2004.11.005

Chen, Z., & Liu, B. (2018). Continual Learning and Catastrophic Forgetting. In *Lifelong Machine Learning* (pp. 55–75). Morgan & Claypool Publishers.

Cheng, L., & Aghakouchak, A. (2014). Nonstationary precipitation intensity-duration-frequency curves for infrastructure design in a changing climate. *Scientific Reports*, *4*(7093). https://doi.org/10.1038/srep07093

Chiang, Y. M., Chang, L. C., Tsai, M. J., Wang, Y. F., & Chang, F. J. (2010). Dynamic neural networks for real-time water level predictions of sewerage systems-covering gauged and ungauged sites. *Hydrology and Earth System Sciences*, *14*(7), 1309–1319. https://doi.org/10.5194/hess-14-1309-2010

Chin, R. J., Lai, S. H., Chang, K. B., Jaafar, W. Z. W., & Othman, F. (2016). Relationship between minimum inter-event time and the number of rainfall

events in Peninsular Malaysia. *Weather*, *71*(9), 213–218. https://doi.org/10.1002/wea.2766

Ciğizoğlu, H. K. (2002). Suspended sediment estimation for rivers using artificial neural networks and sediment rating curves. *Turkish Journal of Engineering and Environmental Sciences*, *26*(1), 27–36.

Civil Engineering Research Foundation. (2001). *Evaluation of SSET: the sewer scanner and evaluation technology technical evaluation report.* ASCE Publications.

CIWEM. (2016). *Event duration monitoring good practice guide*.

Clark, P., Roberts, N., Lean, H., Ballard, S. P., & Charlton-Perez, C. (2016). Convection-permitting models: A step-change in rainfall forecasting. In *Meteorological Applications* (Vol. 23, Issue 2, pp. 165–181). John Wiley and Sons Ltd. https://doi.org/10.1002/met.1538

CRABTREE, R. W. (1989). Sediments in Sewers. *Water and Environment Journal*, *3*(6), 569–578. https://doi.org/10.1111/j.1747-6593.1989.tb01437.x

De Jong, K. (2007). Parameter setting in EAs: A 30 year perspective. In *Studies in Computational Intelligence*. Springer. https://doi.org/10.1007/978-3-540-69432-8_1

Del Mundo, D. M. N., & Sutheerawattananonda, M. (2017). Influence of fat and oil type on the yield, physico-chemical properties, and microstructure of fat, oil, and grease (FOG) deposits. *Water Research*, *124*, 308–319. https://doi.org/10.1016/j.watres.2017.07.047

Detectronic. (2018). *Predicting blockages, reducing flooding and improving network performance*. https://www.detectronic.org/wp-content/uploads/2018/09/MMS-B001-overview_bro.pdf

Detectronic. (2019). *What is a waste water flow meter?* http://wpdev5.thelegaldc.co.uk/what-is-a-waste-water-flow-meter/

Drainage Strategy Framework. (2013). *Good practice guidance commissioned by the Environment Agency and Ofwat*.

Duncan, A. P., Chen, A. S., Keedwell, E. C., Djordjević, S., & Savić, D. A. (2013). RAPIDS: Early warning system for urban flooding and water quality hazards. *Machine Learning in Water Systems - AISB Convention 2013*, 25–29. http://www.scopus.com/inward/record.url?eid=2-s2.0-84894223470&partnerID=tZOtx3y1

Duran, O., Althoefer, K., & Seneviratne, L. D. (2002). State of the art in sensor technologies for sewer inspection. *IEEE Sensors Journal*, *2*(2), 73–81. https://doi.org/10.1109/JSEN.2002.1000245

Ebtehaj, I., & Bonakdari, H. (2014). Performance Evaluation of Adaptive Neural Fuzzy Inference System for Sediment Transport in Sewers. *Water Resources Management*, *28*(13), 4765–4779. https://doi.org/10.1007/s11269-014-0774-0

Edmondson, V., Cerny, M., Lim, M., Gledson, B., Lockley, S., & Woodward, J. (2018). A smart sewer asset information model to enable an 'Internet of Things' for operational wastewater management. *Automation in Construction*, *91*, 193–205. https://doi.org/10.1016/j.autcon.2018.03.003

EEA. (2020). *European Bathing Water Quality in 2019*. https://www.eea.europa.eu//publications/european-bathing-water-quality-in-2019

Ékes, C., Neducza, B., Henrich, G. R., & Corp, S. T. (2011). Gpr Goes Underground : Pipe Penetrating Radar. *Proceedings of North American Society for Trenchless Technology (NASTT), No Dig Show 2011*, 1–10.

EMS. (2020). *SMART Sewer*. https://www.em-solutions.co.uk/innovation/smart-sewer-project/

Environment Agency. (2013). *Good practice guidance commissioned by the Environment Agency and Ofwat* (Issue May).

Environment Agency. (2018a). *Guidance Water companies: environmental permits for storm overflows and emergency overflows*.

Environment Agency. (2018b). *Storm Overflow Assessment Framework - Version 1.6*.

EU UWWTD. (1991). *Urban Wastewater Treatment Directive*. http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:31991L0271&from=E

Evora, D., & Coulibaly, P. (2009). *Recent advances in data-driven modeling of remote sensing applications in hydrology*. 194–201. https://doi.org/10.2166/hydro.2009.036

Fenner, R. A. (2000). Approaches to sewer maintenance: a review. *Urban Water*, *2*(4), 343–356.

Fenner, R. A., McFarland, G., & Thorne, O. (2007). Case-based reasoning approach for managing sewerage assets. *Proceedings of the Institution of Civil Engineers & Water. Management*, *160*(1), 15–24.

Fernando, A. K., Zhang, X., & Kinley, P. F. (2006). Combined Sewer Overflow forecasting with Feed-forward Back-propagation Artificial Neural Network. *Transactions on Engineering, Computing and Technology*, *March*, 58–64.

Fontecha, J. E., Akhavan-Tabatabaei, R., Duque, D., Medaglia, A. L., Torres, M. N., & Rodríguez, J. P. (2016). On the preventive management of sediment-related sewer blockages: A combined maintenance and routing optimization approach. *Water Science and Technology*, *74*(2), 302–308. https://doi.org/10.2166/wst.2016.160

Gaál, L., Molnar, P., & Szolgay, J. (2014). Selection of intense rainfall events based on intensity thresholds and lightning data in Switzerland. *Hydrology and Earth System Sciences*, *18*(5), 1561–1573. https://doi.org/10.5194/hess-18-1561-2014

García, J. T., Espín-Leal, P., Vigueras-Rodriguez, A., Castillo, L. G., Carrillo, J. M., Martínez-Solano, P. D., & Nevado-Santos, S. (2017). Urban runoff characteristics in Combined Sewer Overflows (CSOs): Analysis of storm events in southeastern Spain. *Water*, *9*(5), 303–318. https://doi.org/10.3390/w9050303

Garofalo, G., Giordano, A., Piro, P., Spezzano, G., & Vinci, A. (2017). A distributed real-time approach for mitigating CSO and flooding in urban drainage systems. *Journal of Network and Computer Applications*, *78*, 30–42. https://doi.org/10.1016/j.jnca.2016.11.004

Gelormino, M. S., & Ricker, N. L. (1994). Model-predictive control of a combined sewer system. *International Journal of Control*, *59*(3), 793–816. https://doi.org/10.1080/00207179408923105

Ghiassi, M., Zimbra, D. K., & Saidane, H. (2008). Urban water demand forecasting with a dynamic artificial neural network model. *J. Water Resour. Plann. Manage.*, *134*(2), 138–146. https://doi.org/10.1061

Giusti, R., & Batista, G. E. A. P. A. (2013). An empirical comparison of dissimilarity measures for time series classification. *Proceedings - 2013 Brazilian Conference on Intelligent Systems, BRACIS 2013*, 82–88. https://doi.org/10.1109/BRACIS.2013.22

Giustolisi, O., & Savic, D. A. (2006). A symbolic data-driven technique based on evolutionary polynomial regression. *Journal of Hydroinformatics*, *8*(3), 207–222. https://doi.org/10.2166/hydro.2006.020b

Gnauck, A. (2004). Interpolation and approximation of water quality time series and process identification. *Analytical and Bioanalytical Chemistry*, *380*(3 SPEC.ISS.), 484–492. https://doi.org/10.1007/s00216-004-2799-3

Goh, C.-K., & Tan, K. C. (2009). Evolutionary Multi-objective Optimization in Uncertain Environments. In *Evolutionary Multi-objective Optimization in Uncertain Environments*. Springer. https://doi.org/10.1007/978-3-540-95976-2

Golding, B. W., Ballard, S. P., Mylne, K., Roberts, N., Saulter, A., Wilson, C., Agnew, P., Davis, L. S., Trice, J., Jones, C., Simonin, D., Li, Z., Pierce, C., Bennett, A., Weeks, M., & Moseley, S. (2014). Forecasting capabilities for the london 2012 olympics. *Bulletin of the American Meteorological Society*, *95*(6), 883–896. https://doi.org/10.1175/BAMS-D-13-00102.1

Grandison, J. (2005). Sewer Asset Management Through Long Term CSO Monitoring. *WaPUG Spring Conference*.

Gu, Y., Wang, H. O., Tanaka, K., & Bushnell, L. G. (2001). Fuzzy control of nonlinear time-delay systems: Stability and design issues. *Proceedings of the American Control Conference*, *6*, 4771–4776. https://doi.org/10.1109/ACC.2001.945736

Guo, N., & Saul, A. J. (2011). Improving the operation and maintenance of CSO structures. *12th International Conference on Urban Drainage*, 1–8.

Hafskjold, L. S., König, A., Sægrov, S., & Schilling, W. (2004). Improved assessment of sewer pipe condition. *CityNet 19th European Junior Scientist Workshop, Lyon, 11–14 March 2004*.

Halfawy, M. R., & Hengmeechai, J. (2014). Automated defect detection in sewer closed circuit television images using histograms of oriented gradients and support vector machine. *Automation in Construction*, *38*, 1–13. https://doi.org/10.1016/j.autcon.2013.10.012

He, H., & Ma, Y. (Eds.). (2013). *Imbalanced Learning: Foundations, Algorithms, and Applications* (1st ed.). Wiley-IEEE Press.

Hillas, T. T. (2014). Reducing the Occurrence of Flooding through the Effective Management of Sewer Blockages. *PhD Thesis, University of Exeter.*

Hines, A., Skoglund, J., Kokaram, A., & Harte, N. (2012). ViSQOL: the virtual speech quality objective listener. *Proceedings of the International Workshop*

on Acoustic Signal Enhancement, Aachen, Germany, 1–4.

Hines, Andrew, & Harte, N. (2012). Speech intelligibility prediction using a Neurogram Similarity Index Measure. *Speech Communication*, *54*(2), 306–320. https://doi.org/10.1016/j.specom.2011.09.004

Holland, J. H. (1984). Genetic Algorithms and Adaptation. In *Adaptive Control of Ill-Defined Systems* (pp. 317–333). Springer US. https://doi.org/10.1007/978-1-4684-8941-5_21

Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, *2*(5), 359–366. https://doi.org/10.1016/0893-6080(89)90020-8

House-Peters, L. A., & Chang, H. (2011). Urban water demand modeling: Review of concepts, methods, and organizing principles. *Water Resources Research*, *47*. https://doi.org/10.1029/2010WR009624

Hu, Z., Chen, B., Chen, W., Tan, D., & Shen, D. (2021). Review of model-based and data-driven approaches for leak detection and location in water distribution systems. *Water Supply*. https://doi.org/10.2166/ws.2021.101

Huynh, P., Ross, R., Martchenko, A., & Devlin, J. (2016). Anomaly inspection in sewer pipes using stereo vision. *IEEE 2015 International Conference on Signal and Image Processing Applications, ICSIPA 2015 - Proceedings*, 60–64. https://doi.org/10.1109/ICSIPA.2015.7412164

Hwang, S. H., Ham, D. H., & Kim, J. H. (2012). A new measure for assessing the efficiency of hydrological data-driven forecasting models. *Hydrological Sciences Journal*, *57*(7), 1257–1274. https://doi.org/10.1080/02626667.2012.710335

IBAK. (n.d.). *IBAK Helmut Hunger GmbH & Co. KG: Products*. Retrieved August 15, 2020, from https://www.ibak.de/en/produkte/ibak_show/frontendshow/category/

Innovyze. (n.d.-a). *InfoSewer* ®. Retrieved August 18, 2020, from www.innovyze.com

Innovyze. (n.d.-b). *InfoSWMM Real-Time Control (RTC) Rule Editor* . Retrieved August 18, 2020, from https://swmm5.org/2018/11/10/infoswmm-real-time-control-rtc-rule-editor-2019/

Institution of Public Health Engineers. (1954). *Report of the Joint Committee on Field Research into Drainage Problems*.

Institution of Water and Environmental Management (IWEM). (1993). *Glossary. Handbooks of UK Wastewater Practice*.

Isermann, R. (1984). Process fault detection based on modeling and estimation methods-A survey. *Automatica*, *20*(4), 387–404. https://doi.org/10.1016/0005-1098(84)90098-0

Islam, M. M. M., Hofstra, N., & Islam, M. A. (2017). The Impact of Environmental Variables on Faecal Indicator Bacteria in the Betna River Basin, Bangladesh. *Environmental Processes*, *4*(2), 319–332. https://doi.org/10.1007/s40710-017-0239-6

Izquierdo, J., Pérez, R., & Iglesias, P. L. (2004). Mathematical models and methods in the water industry. *Mathematical and Computer Modelling*,

39(11–12), 1353–1374. https://doi.org/10.1016/j.mcm.2004.06.012

Jafari, S., & Mashohor, S. (2011). Robust combining methods in committee neural networks. *Proceedings of 2011 IEEE Symposium on Computers and Informatics (ISCI)*.

Jalliffier-Verne, I., Heniche, M., Madoux-Humery, A. S., Galarneau, M., Servais, P., Prévost, M., & Dorner, S. (2016). Cumulative effects of fecal contamination from combined sewer overflows: Management for source water protection. *Journal of Environmental Management*, *174*, 62–70. https://doi.org/10.1016/j.jenvman.2016.03.002

James, W. (1994). *Current Practices in Modelling the Management of Stormwater Impacts* (1st ed.). Taylor and Francis Ltd.

Jin, Y., & Mukherjee, A. (2010). Modeling Blockage Failures in Sewer Systems to Support Maintenance Decision Making. *Journal of Performance of Constructed Facilities*, *24*(6), 622–633. https://doi.org/10.1061/(asce)cf.1943-5509.0000126

Joo, J., Lee, J., Kim, J. H., Jun, H., & Jo, D. (2014). Inter-event time definition setting procedure for urban drainage systems. *Water*, *6*(1), 45–58. https://doi.org/10.3390/w6010045

Joseph-Duran, B., Jung, M. N., Ocampo-Martinez, C., Sager, S., & Cembrano, G. (2014). Minimization of Sewage Network Overflow. *Water Resources Management*, *28*(1), 41–63. https://doi.org/10.1007/s11269-013-0468-z

Karl, T. R., & Knight, R. W. (1998). Secular Trends of Precipitation Amount, Frequency, and Intensity in the United States. *Bulletin of the American Meteorological Society*, *79*(2), 231–241. https://doi.org/10.1175/1520-0477(1998)079<0231:STOPAF>2.0.CO;2

Katipamula, S., & Brambley, M. (2005a). Review Article: Methods for Fault Detection, Diagnostics, and Prognostics for Building Systems—A Review, Part II. *HVAC&R Research*, *11*(2), 169–187. https://doi.org/10.1080/10789669.2005.10391133

Katipamula, S., & Brambley, M. R. (2005b). Review article: Methods for fault detection, diagnostics, and prognostics for building systems—A review, part I. *HVAC and R Research*, *11*(1), 3–25. https://doi.org/10.1080/10789669.2005.10391123

Kendall, M. G. (1975). *Rank Correlation Methods.* (4th ed.). Griffin.

Khan, M. S., & Patil, R. (2018). Acoustic Characterization of PVC Sewer Pipes for Crack Detection Using Frequency Domain Analysis. *2018 IEEE International Smart Cities Conference, ISC2 2018*. https://doi.org/10.1109/ISC2.2018.8656739

Kim, H. Il, & Han, K. Y. (2020). Urban flood prediction using deep neural network with data augmentation. *Water*, *12*(3). https://doi.org/10.3390/w12030899

Kisi, O. (2005). Suspended sediment estimation using neuro-fuzzy and neural network approaches. *Hydrological Sciences Journal*, *50*(4), 683–696. https://doi.org/10.1623/hysj.2005.50.4.683

Korving, H., Clemens, F. H. L. R., & van Noortwijk, J. M. (2006). Statistical modeling of the serviceability of sewage pumps. *Journal of Hydraulic Engineering*, *132*(10), 1076–1085. https://doi.org/10.1061/(ASCE)0733-

9429(2006)132:10(1076)

Koutsoyiannis, D., Kozonis, D., & Manetas, A. (1998). A mathematical framework for studying rainfall intensity-duration-frequency relationships. *Journal of Hydrology*, *206*(1–2), 118–135. https://doi.org/10.1016/S0022-1694(98)00097-3

Koza, J. R. (1994). Genetic programming as a means for programming computers by natural selection. *Statistics and Computing*, *4*(2), 87–112. https://doi.org/10.1007/BF00175355

Kuliczkowska, E. (2008). *Kryteria planowania bezwykopowej odnowy nie przełazowych przewodów kanalizacyjnych (Criteria of planning of the trenchless renewal of non entry sewers)*. Kielce University of Technology, Kielce.

Kumar, S. S., Wang, M., Abraham, D. M., Jahanshahi, M. R., Iseley, T., & Cheng, J. C. P. (2020). Deep Learning–Based Automated Detection of Sewer Defects in CCTV Videos. *Journal of Computing in Civil Engineering*, *34*(1), 04019047. https://doi.org/10.1061/(asce)cp.1943-5487.0000866

Kurth, A., Saul, A., Mounce, S., Shepherd, W., & Hanson, D. (2008). Application of Artificial Neural Networks (ANNs) for the prediction of CSO discharges. *Proceedings of 11th International Conference on Urban Drainage*, 1–10.

Launay, M. A., Dittmer, U., & Steinmetz, H. (2016). Organic micropollutants discharged by combined sewer overflows – Characterisation of pollutant sources and stormwater-related processes. *Water Research*, *104*, 82–92. https://doi.org/10.1016/j.watres.2016.07.068

Lee, D.-H., & Kang, D.-S. (2016). The Application of the Artificial Neural Network Ensemble Model for Simulating Streamflow. *12th International Conference on Hydroinformatics*, 1217 – 1224.

Lee, E. H., & Kim, J. H. (2018). Development of New Inter-Event Time Definition Technique in Urban Areas. *KSCE Journal of Civil Engineering*, *22*(10), 3764–3771. https://doi.org/10.1007/s12205-018-1120-5

Lee, L. J. E., Lawrence, D. S. L., & Price, M. (2006). Analysis of water-level response to rainfall and implications for recharge pathways in the Chalk aquifer, SE England. *Journal of Hydrology*, *330*(3–4), 604–620. https://doi.org/10.1016/j.jhydrol.2006.04.025

Legates, D. R., & McCabe, G. J. (1999). Evaluating the use of “goodness-of-fit” measures in hydrologic and hydroclimatic model validation. *Water Resources Research*, *35*(1), 233–241. https://doi.org/10.1029/1998WR900018

Lepot, M., Aubin, J. B., & Clemens, F. H. L. R. (2017). Interpolation in time series: An introductive overview of existing methods, their performance criteria and uncertainty assessment. *Water*, *9*(10). https://doi.org/10.3390/w9100796

Levandowsky, M., & Winter, D. (1971). Distance between sets. In *Nature* (Vol. 234, Issue 5323, pp. 34–35). Nature Publishing Group. https://doi.org/10.1038/234034a0

Li, J. (2020). A data-driven improved fuzzy logic control optimization-simulation tool for reducing flooding volume at downstream urban drainage systems. *Science of the Total Environment*, *732*, 138931.

https://doi.org/10.1016/j.scitotenv.2020.138931

Li, X., Zhou, F., & Lodewyk, S. (2010). Applications of Artificial Neural Networks in Urban Water System. *Watershed Management Conference, Madison, Wis*.

Liguori, S., Rico-Ramirez, M. A., Schellart, A. N. A., & Saul, A. J. (2012). Using probabilistic radar rainfall nowcasts and NWP forecasts for flow prediction in urban catchments. *Atmospheric Research*, *103*, 80–95. https://doi.org/10.1016/j.atmosres.2011.05.004

Lillywhite, M. S. T., & Webster, C. J. D. (1979). Investigations of Drain Blockages and Their Implications on Design. *The Public Health Engineer*, *7*, 53–60.

Lincoln, W. P., & Skrzypek, J. (1989). Synergy of clustering multiple back propagation networks. *Advances in Neural Information Processing Systems 2 (NIPS 1989)*, 650–657.

Liu, X. (2012). Classification accuracy and cut pointselection. *Statistics in Medicine*, *31*(23), 2676–2686. https://doi.org/10.1002/sim.4509

Liu, Z., Dai, Q., & Zhuo, L. (2019). Relationship between Rainfall Variability and the Predictability of Radar Rainfall Nowcasting Models. *Atmosphere*, *10*(8), 458. https://doi.org/10.3390/atmos10080458

López-Vázquez, G., Espinal, A., Ornelas-Rodríguez, M., Soria-Alcaraz, J. A., Rojas-Domínguez, A., Puga, H., Carpio, J. M., & Rostro-González, H. (2020). Comparing Evolutionary Artificial Neural Networks from Second and Third Generations for Solving Supervised Classification Problems. In *Studies in Computational Intelligence* (Vol. 862, pp. 615–628). Springer. https://doi.org/10.1007/978-3-030-35445-9_42

Lund, N. S. V., Falk, A. K. V., Borup, M., Madsen, H., & Steen Mikkelsen, P. (2018). Model predictive control of urban drainage systems: A review and perspective towards smart real-time water management. *Critical Reviews in Environmental Science and Technology*, *48*(3), 279–339. https://doi.org/10.1080/10643389.2018.1455484

Ma, H., Bandos, A. I., Rockette, H. E., & Gur, D. (2013). On use of partial area under the ROC curve for evaluation of diagnostic performance. *Statistics in Medicine*, *32*(20), 3449–3458. https://doi.org/10.1002/sim.5777

Mackay, J. D., Jackson, C. R., & Wang, L. (2014). A lumped conceptual model to simulate groundwater level time-series. *Environmental Modelling and Software*, *61*, 229–245. https://doi.org/10.1016/j.envsoft.2014.06.003

Maier, H. R., & Dandy, G. C. (2001). Neural network based modelling of environmental variables: A systematic approach. *Mathematical and Computer Modelling*, *33*(6–7), 669–682. https://doi.org/10.1016/S0895-7177(00)00271-5

Marinaki, M., & Papageorgiou, M. (1998). Nonlinear optimal flow control for sewer networks. *Proceedings of the American Control Conference, Philadelphia, Pennsylvania.*, *2*, 1289–1293.

Marlow, D. R., Boulaire, F., Beale, D. J., Grundy, C., & Moglia, M. (2011). *Sewer Performance Reporting : Factors That Influence Blockages*. *17*(1), 42–51. https://doi.org/10.1061/(ASCE)IS.1943-555X.0000041.

Martin, S. (2017). Tackling the impact of FOG in drainage systems. *Filtration and*

Separation, *54*(2), 42–44. https://doi.org/10.1016/S0015-1882(17)30086-1

Massari, C., Yeh, T. J., Ferrante, M., Brunone, B., & Meniconi, S. (2014). A stochastic tool for determining the presence of partial blockages in viscoelastic pipelines : first experimental results. *Procedia Engineering*, *70*, 1112–1120. https://doi.org/10.1016/j.proeng.2014.02.123

May, R. W. . (1981). Sediment transport in rivers. *Hydraulic Research Station (Wallingford), Report IT 222.* https://doi.org/10.1201/noe0415453639-p9

McClelland, J., Rumelhart, D., & PDP Research Group. (1986). Parallel distributed processing: Explorations in the microstructure of cognition. In *Psychological and Biological Models*. The MIT Press.

McCoy, K. J., & Blanchard, P. J. (2008). Precipitation, ground-water hydrology and recharge along the eastern slopes of the Sandia Mountains, Bernalillo County, New Mexico. *U.S. Geological Survey Scientific Investigations Report 2008-5179*, 34. http://www.usgs.gov/

McCuen, R. H., Knight, Z., & Cutter, A. G. (2006). Evaluation of the Nash–Sutcliffe Efficiency Index. *Journal of Hydrologic Engineering*, *11*(6), 597–602. https://doi.org/10.1061/(asce)1084-0699(2006)11:6(597)

Mcintyre, N., Dı, M., & Rodrı, J. P. (2012). *A database and model to support proactive management of sediment-related sewer blockages*. *6*(0). https://doi.org/10.1016/j.watres.2012.06.037

Meniconi, S., Brunone, B., & Ferrante, M. (2009). In-line partially closed valves: How to detect by transient tests. *World Environmental and Water Resources Congress 2009*, *342*, 135–144. https://doi.org/10.1061/41036(342)14

Meniscus. (n.d.). *Sewer level monitoring, CSO monitoring, sewer level analytics*. Retrieved August 18, 2020, from http://www.meniscus.co.uk/meniscus-analytics-case-studies/analytics-for-the-water-industry/real-time-sewer-and-cso-monitoring/

Met Office. (2007). *Fact sheet No. 3 – Water in the atmosphere*.

Met Office. (2009). *Fact Sheet 15, Weather Radar*. http://www.metoffice.gov.uk/media/pdf/o/c/fact_sheet_No. _15.pdf

Metz, C. E. (1978). Basic principles of ROC analysis. *Seminars in Nuclear Medicine*, *8*(4), 283-98.

Mo, R., Ye, C., & Whitfield, P. (2013). *Some Similarity Indices with Potential Meteorological Applications*. http://www3.telus.net/ruping/lib/reports/2013-002-CMML-TR_Mo_etal.pdf

Moore, R. J. (2014). Weather Radar and Hydrology: Preface. *Hydrological Sciences Journal*, *59*(7), 1275. https://doi.org/10.1080/02626667.2014.934080

Moradi, A. M., & Dariane, A. B. (2017). Evolving neural networks and fuzzy clustering for multireservoir operations. *Neural Computing and Applications*, *28*(5), 1149–1162. https://doi.org/10.1007/s00521-015-2130-6

Morales, V. M., Mier, J. M., & Garcia, M. H. (2017). Innovative modeling framework for combined sewer overflows prediction. *Urban Water Journal*, *14*(1), 97–111. https://doi.org/10.1080/1573062X.2015.1057183

Mounce, S. R., Boxall, J. B., & Machell, J. (2010). Development and Verification

of an Online Artificial Intelligence System for Detection of Bursts and Other Abnormal Flows. *Journal of Water Resources Planning and Management*, *136*(3), 309–318. https://doi.org/10.1061/(asce)wr.1943-5452.0000030

Mounce, S., Shepherd, W., Ostojin, S., Abdel-Aal, M., Schellart, A. N. A., Shucksmith, J. D., & Tait, S. J. (2020). Optimisation of a fuzzy logic-based local real-time control system for mitigation of sewer flooding using genetic algorithms. *Journal of Hydroinformatics*, *22*(2), 281–295. https://doi.org/10.2166/hydro.2019.058

Mounce, S., Shepherd, W., Sailor, G., Saul, A., & Boxhall, J. (2014a). Application of Artificial Neural Networks to Assess CSO Performance. *13th International Conference on Urban Drainage*.

Mounce, S., Shepherd, W., Sailor, G., Shucksmith, J., & Saul, A. J. (2014b). Predicting combined sewer overflows chamber depth using artificial neural networks with rainfall radar data. *Water Science and Technology*, *69*(6), 1326–1333. https://doi.org/10.2166/wst.2014.024

Murali, M. K., Hipsey, M. R., Ghadouani, A., & Yuan, Z. (2020). SewerSedFoam: A Model for Free Surface Flow, Sediment Transport, and Deposited Bed Morphology in Sewers. *Water*, *12*(1), 270. https://doi.org/10.3390/w12010270

Murali, M. K., Hipsey, M. R., Ghadouani, A., & Yuan, Z. (2019). Developing and Validating a Model to Assess Sewer Sediment Issues from Changing Wastewater Inflows and Concentration. *Green Energy and Technology*, 836–841. https://doi.org/10.1007/978-3-319-99867-1_144

Nalluri, C., El-Zaemey, A. K., & Chan, H. L. (1997). Sediment transport over fixed deposited beds in sewers - An appraisal of existing models. *Water Science and Technology*, *36*(8–9), 123–128. https://doi.org/10.1016/S0273-1223(97)00609-4

Nalluri, C., Ghani Ab., A., & El-Zaemey, A. K. S. (1994). Sediment transport over deposited beds in sewers. *Water Science and Technology*, *29*(1–2), 125–133. https://doi.org/10.2166/wst.1994.0658

National Geographic. (2017). *Huge Blobs of Fat and Trash Are Filling the World's Sewers*. National Geographic. https://www.nationalgeographic.com/news/2017/08/fatbergs-fat-cities-sewers-wet-wipes-science/

Nicolau, A. D. S., Augusto, J. P. D. S. C., & Schirru, R. (2017). Accident diagnosis system based on real-time decision tree expert system. *AIP Conference Proceedings*, *1836*. https://doi.org/10.1063/1.4981957

Norazian, M. N., Shukri, Y. A., Azam, R. N., & Al Bakri, A. M. M. (2008). Estimation of missing values in air pollution data using single imputation techniques. *ScienceAsia*, *34*(3), 341–345. https://doi.org/10.2306/scienceasia1513-1874.2008.34.341

Nuron. (n.d.-a). *Nuron fibre sensing technology*. Retrieved August 22, 2020, from http://www.nuron.tech/technology/

Nuron. (n.d.-b). *World's first sewer nervous system implementation is up and running*. Retrieved August 22, 2020, from http://www.nuron.tech/news/worlds-first-sewer-nervous-system-implementation-is-up-and-running/

Ocampo-Martínez, C., Puig, V., Quevedo, J., & Ingimundarson, A. (2005). Fault tolerant model predictive control applied on the Barcelona sewer network. *Proceedings of the 44th IEEE Conference on Decision and Control, and the European Control Conference, CDC-ECC '05*, *2005*, 1349–1354. https://doi.org/10.1109/CDC.2005.1582346

Odan, F. K., & Reis, L. F. R. (2012). Hybrid Water Demand Forecasting Model Associating Artificial Neural Network with Fourier Series. *Journal of Water Resources Planning and Management*, *138*(3), 245–256. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000177

Ofwat. (2017a). *Appendix 2: Delivering outcomes for customers Appendix to Chapter 4: Delivering outcomes for customers*.

Ofwat. (2017b). *Delivering Water 2020: Our final methodology for the 2019 price review*. https://www.ofwat.gov.uk/wp-content/uploads/2017/12/Final-methodology-1.pdf

Ofwat. (2017c). *Outcomes definitions - PR19 -*. https://www.ofwat.gov.uk/outcomes-definitions-pr19/

Ofwat. (2019). *PR19 final determinations: Severn Trent Water - Outcomes performance commitment appendix*. https://www.ofwat.gov.uk/wp-content/uploads/2019/12/PR19-final-determinations-Severn-Trent-Water-Outcomes-performance-commitment-appendix.pdf

Opitz, D. W., & Shavlik, J. W. (1996). Actively Searching for an Effective Neural Network Ensemble. *Connection Science*, *8*(3–4), 337–354. https://doi.org/10.1080/095400996116802

Osborn, T. J., Hulme, M., Jones, P. D., & Basnett, T. A. (2000). Observed trends in the daily intensity of United Kingdom precipitation. *International Journal of Climatology*, *20*(4), 347–364. https://doi.org/10.1002/(SICI)1097-0088(20000330)20:4<347::AID-JOC475>3.0.CO;2-C

Page, E. S. (1955). Control Charts with Warning Lines. *Biometrika*, *42*(1/2), 243. https://doi.org/10.2307/2333440

Peng, J., Wang, H., Li, J., & Gao, H. (2016). Set-based similarity search for time series. *Proceedings of the ACM SIGMOD International Conference on Management of Data*, *26-June-2016*, 2039–2052. https://doi.org/10.1145/2882903.2882963

Perkins, N. J., & Schisterman, E. F. (2006). The Inconsistency of " Optimal " Cut-points Using Two ROC Based Criteria. *American Journal of Epidemiology*, *163*(7), 670–675. https://doi.org/10.1093/aje/kwj063

Phillips, P. J., Chalmers, A. T., Gray, J. L., Kolpin, D. W., Foreman, W. T., & Wall, G. R. (2012). Combined Sewer Overflows: An Environmental Source of Hormones and Wastewater Micropollutants. *Environ. Sci. Technol*, *46*(10), 52. https://doi.org/10.1021/es3001294

Piotrowski, A. P., & Napiorkowski, J. J. (2013). A comparison of methods to avoid overfitting in neural networks training in the case of catchment runoff modelling. *Journal of Hydrology*, *476*, 97–111. https://doi.org/10.1016/j.jhydrol.2012.10.019

Polade, S. D., Pierce, D. W., Cayan, D. R., Gershunov, A., & Dettinger, M. D. (2014). The key role of dry days in changing regional climate and

precipitation regimes. *Scientific Reports*, *4*. https://doi.org/10.1038/srep04364

Prawira Putra, Y., Khrisne, D. C., & Suyadnya, I. M. A. (2019). Expert System for Early Diagnosis of Heart Disease Using the Random Forest Method. *Journal of Electrical, Electronics and Informatics*, *3*(1), 15. https://doi.org/10.24843/jeei.2019.v03.i01.p03

Puig, V., Cembrano, G., Romera, J., Quevedo, J., Aznar, B., Ramón, G., & Cabot, J. (2009). Predictive optimal control of sewer networks using CORAL tool: Application to Riera Blanca catchment in Barcelona. *Water Science and Technology*, *60*(4), 869–878. https://doi.org/10.2166/wst.2009.424

Pulido, E. S., Arboleda, C. V., & Rodríguez Sánchez, J. P. (2019). Study of the spatiotemporal correlation between sediment-related blockage events in the sewer system in Bogotá (Colombia). *Water Science and Technology*, *79*(9), 1727–1738. https://doi.org/10.2166/wst.2019.172

Qian, W., Fu, J., & Yan, Z. (2007). Decrease of light rain events in summer associated with a warming environment in China during 1961-2005. *Geophysical Research Letters*, *34*(11), 179–195. https://doi.org/10.1029/2007GL029631

RedZone Robotics. (n.d.). *Solo Rapid Autonomous CCTV Inspection by RedZone Robotics*. Retrieved August 15, 2020, from https://redzone.com/technology/inspection-equipment/solo/

Regneri, M., Klepiszewski, K., Muschalla, D., & Vanrolleghem, P. A. (2015). Integrated multi-criteria optimal model predictive control of a sewer network in a rural catchment. *10th International Urban Drainage Modelling Conference*.

Remesan, R., & Mathew, J. (2015). Hydrological data driven modelling: A case study approach. In *Hydrological Data Driven Modelling: A Case Study Approach*. Springer International Publishing. https://doi.org/10.1007/978-3-319-09235-5

Roberts, S. (1958). Properties of control chart zone tests. *Bell System Technical Journal*, *37*(1), 83–114. https://doi.org/https://doi.org/10.1002/j.1538-7305.1958.tb03870.x

Romano, M., & Kapelan, Z. (2014a). Adaptive water demand forecasting for near real-time management of smart water distribution systems. *Environmental Modelling and Software*, *60*, 265–276. https://doi.org/10.1016/j.envsoft.2014.06.016

Romano, M., & Kapelan, Z. (2014b). Adaptive water demand forecasting for near real-time management of smart water distribution systems. *Environmental Modelling and Software*, *60*, 265–276. https://doi.org/10.1016/j.envsoft.2014.06.016

Romano, M., Kapelan, Z., & Savic, D. A. (2012). Testing of the system for detection of pipe bursts and other events in a UK water distribution system. *14th Water Distribution Systems Analysis Conference 2012, WDSA 2012*, *2*, 1109–1126.

Romano, M., Kapelan, Z., & Savić, D. A. (2013). Geostatistical techniques for approximate location of pipe burst events in water distribution systems. *Journal of Hydroinformatics*, *15*(3), 634–651.

https://doi.org/10.2166/hydro.2013.094

Rosin, P. L., & Fierens, F. (1995). Improving neural network generalisation. *International Geoscience and Remote Sensing Symposium (IGARSS)*, *2*, 1255–1257. https://doi.org/10.1109/igarss.1995.521718

Saenz, J., Elkmann, N., Stuerze, T., Kutzner, S., & Althoff, H. (2010). Robotic systems for cleaning and inspection of large concrete pipes. *St International Conference on Applied Robotics for the Power Industry, CARPI 2010*. https://doi.org/10.1109/CARPI.2010.5624448

Salerno, F., Viviano, G., & Tartari, G. (2018). Urbanization and climate change impacts on surface water quality: Enhancing the resilience by reducing impervious surfaces. *Water Research*, *144*, 491–502. https://doi.org/10.1016/j.watres.2018.07.058

Savić, D. A., Giustolisi, O., Berardi, L., Shepherd, W., Djordjević, S., & Saul, A. (2006). Modelling sewer failure by evolutionary computing. *Proceedings of the ICE - Water Management*, *159*(June), 111–118. https://doi.org/10.1680/wama.2006.159.2.111

Schladweiler, J. C. (2004). *Tracking Down the Roots of Our Sanitary Sewers*. Arizona Water & Pollution Control Association.

Schwefel, H.-P. (1998). Evolution and optimum seeking. In *Kybernetes* (Vol. 27, Issue 8). Wiley. https://doi.org/10.1108/k.1998.27.8.975.2

Scott, D. (1979). On optimal and data-based histograms. *Biometrika*, *66*, 605–610.

Shahra, E. Q., Wu, W., & Romano, M. (2019). Considerations on the deployment of heterogeneous IoT devices for smart water networks. *The 16 IEEE International Conference on Ubiquitous Intelligence and Computing (UIC 2019)*, 791–796. https://doi.org/10.1109/SmartWorld-UIC-ATC-SCALCOM-IOP-SCI.2019.00167

Sharif, H. O., Yates, D., Roberts, R., & Mueller, C. (2006). The use of an automated nowcasting system to forecast flash floods in an urban watershed. *Journal of Hydrometeorology*, *7*(1), 190–202. https://doi.org/10.1175/JHM482.1

Sharma, T., Tiwari, N., & Kelkar, D. (2012). Study of Difference Between Forward and Backward Reasoning. *International Journal of Emerging Technology and Advanced Engineering Website: Www.Ijetae.Com*, *2*(10), 271–273.

She, L., & You, X. yi. (2019). A Dynamic Flow Forecast Model for Urban Drainage Using the Coupled Artificial Neural Network. *Water Resources Management*, *33*(9), 3143–3153. https://doi.org/10.1007/s11269-019-02294-9

Sheela, K. G., & Deepa, S. N. (2013). Review on methods to fix number of hidden neurons in neural networks. *Mathematical Problems in Engineering*, *2013*. https://doi.org/10.1155/2013/425740

Simonin, D., Pierce, C., Roberts, N., Ballard, S. P., & Li, Z. (2017). Performance of Met Office hourly cycling NWP-based nowcasting for precipitation forecasts. *Quarterly Journal of the Royal Meteorological Society*, *143*(708), 2862–2873. https://doi.org/10.1002/qj.3136

Skipworth, P. J., Tait, S. J., & Saul, A. J. (2000). The first foul flush in combined sewers: An investigation of the causes. *Urban Water*, *2*(4), 317–325.

https://doi.org/10.1016/s1462-0758(01)00009-7

Smits, N. (2010). A note on Youden's J and its cost ratio. *BMC Medical Research Methodology Volume*, *10*(89).

Sohn, B. J., Ryu, G. H., Song, H. J., & Ou, M. L. (2013). Characteristic features of warm-type rain producing heavy rainfall over the korean peninsula inferred from TRMM measurements. *Monthly Weather Review*, *141*(11), 3873–3888. https://doi.org/10.1175/MWR-D-13-00075.1

Solomatine, D., See, L. M., & Abrahart, R. J. (2008). Data-Driven Modelling: Concepts, Approaches and Experiences. In *Practical Hydroinformatics* (pp. 17–30). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-79881-1_2

Southern Water. (2018). *Making fatbergs a thing of the past – SMART Sewer*. https://www.southernwater.co.uk/the-news-room/the-media-centre/2018/september/making-fatbergs-a-thing-of-the-past-smart-sewer

Stanić, N., Lepot, M., Catieau, M., Langeveld, J., & Clemens, F. H. L. R. (2017). A technology for sewer pipe inspection (part 1): Design, calibration, corrections and potential application of a laser profiler. *Automation in Construction*, *75*, 91–107. https://doi.org/10.1016/j.autcon.2016.12.005

Storn, R., & Price, K. (1997). Differential Evolution - A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces. *Journal of Global Optimization*, *11*(4), 341–359. https://doi.org/10.1023/A:1008202821328

Stovin, V. R., & Saul, A. J. (2000). Computational Fluid Dynamics and the Design of Sewage Storage Chambers. *Water and Environment Journal*, *14*(2), 103–110. https://doi.org/10.1111/j.1747-6593.2000.tb00235.x

Streich, H., & Adria, O. (2004). Software approach for the autonomous inspection robot MAKRO. *Proceedings - IEEE International Conference on Robotics and Automation*, *2004*(4), 3411–3416. https://doi.org/10.1109/robot.2004.1308781

Svensen, J. L., Niemann, H. H., & Poulsen, N. K. (2019). Model Predictive Control of Overflow in Sewer Networks: A comparison of two methods. *Conference on Control and Fault-Tolerant Systems, SysTol*, 412–417. https://doi.org/10.1109/SYSTOL.2019.8864755

Tadeusiewicz, R. (1995). Neural networks: A comprehensive foundation. In *Control Engineering Practice* (Vol. 3, Issue 5). Pearson Prentice Hall. https://doi.org/10.1016/0967-0661(95)90080-2

Technolog. (n.d.). *Cello CSO - Waste water monitoring with ultrasonic level recorder*. Retrieved July 7, 2020, from https://www.technolog.com/products/cello-cso-gsm-ultrasonic-level-recorder/

Ten Veldhuis, J. A. E., & Clemens, F. H. L. R. (2011). The efficiency of asset management strategies to reduce urban flood risk. *Water Science and Technology*, *64*(6), 1317–1324. https://doi.org/10.2166/wst.2011.715

The New York Times. (2019). *210-Foot Fatberg Blocks Sewers of English Seaside Town*. The New York Times. https://www.nytimes.com/2019/01/08/world/europe/uk-fatberg-sidmouth.html

Thomson, J. C., & Grada, L. (2004). An Examination of Innovative Methods Used in the Inspection of Wastewater Collection Systems. In *Werf Report: Collection Systems 01-cts-7*. IWA Publishing. https://doi.org/10.2166/9781780404349

Thorndahl, S., & Rasmussen, M. R. (2013). Short-term forecasting of urban storm water runoff in real-time using extrapolated radar rainfall data. *Journal of Hydroinformatics*, *15*(3), 897–912. https://doi.org/10.2166/hydro.2013.161

Tirogo, J., Jost, A., Biaou, A., Valdes-Lao, D., Koussoubé, Y., & Ribstein, P. (2016). Climate Variability and Groundwater Response: A Case Study in Burkina Faso (West Africa). *Water*, *8*(5), 171. https://doi.org/10.3390/w8050171

Tripathy, A., & Schwefel, H.-P. (1982). Numerical Optimization of Computer Models. *The Journal of the Operational Research Society*, *33*(12), 1166. https://doi.org/10.2307/2581158

Tuccillo, M., E., Jolley, J., Martel, K., & Boyd, G. (2009). Condition Assessment of Wastewater Collection Systems White Paper. In *United States Environmental Protection Agency*.

Ugarelli, R., Kristensen, S. M., Røstum, J., Sægrov, S., & Di Federico, V. (2009). Statistical analysis and definition of blockages-prediction formulae for the wastewater network of Oslo by evolutionary computing. *Water Science and Technology*, *59*(8), 1457–1470. https://doi.org/10.2166/wst.2009.152

Ugarelli, R., Venkatesh, G., Brattebø, H., Di Federico, V., & Sægrov, S. (2010). Historical analysis of blockages in wastewater pipelines in Oslo and diagnosis of causative pipeline characteristics. *Urban Water Journal*, *7*(6), 335–343. https://doi.org/10.1080/1573062X.2010.526229

UKWIR. (2010). *Economic Level of Service for Sewer Blockages*.

Unal, I. (2017). Defining an optimal cut-point value in ROC analysis: An alternative approach. *Computational and Mathematical Methods in Medicine*, *2017*. https://doi.org/10.1155/2017/3762651

Veldhuis, J. A. E. ten, Clemens, F. H. L. R., Sterk, G., & Berends, B. R. (2010). Microbial risks associated with exposure to pathogens in contaminated urban flood water. *Water Research, 44*(9), 2910–2918.

Venkatasubramanian, V., & Rengaswamy, R. (2003). A review of process fault detection and diagnosis Part I : Quantitative model-based methods. *Computers & Chemical …*, *27*, 293–311. https://doi.org/10.1016/S0098-1354(02)00161-8

Vluymans, S. (2019). Learning from imbalanced data. *Studies in Computational Intelligence*, *807*(9), 81–110. https://doi.org/10.1007/978-3-030-04663-7_4

Volna, E. (2009). Evolutionary artificial neural networks. *Mendel*, *367*(1), 171–176. https://doi.org/10.1109/ccece.1997.614852

Wallerstein, W., & Arthur, S. (2012). Predicting and managing flood risk associated with trash screens at culverts. *FRMRC 2 WP4.1 Final Science Report*.

Wang, X. J., Lambert, M. F., & Simpson, A. R. (2005). Detection and Location of a Partial Blockage in a Pipeline Using Damping of Fluid Transients. *Journal of Water Resources Planning and Management*, *131*(3).

https://doi.org/10.1061/(ASCE)0733-9496(2005)131:3(244)

Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, *13*(4), 600–612. https://doi.org/10.1109/TIP.2003.819861

Water UK. (2017). *Wipes in sewer blockage study – Final Report*. file:///C:/Users/tr319/Downloads/Wipes in Sewer Blockage Study.pdf

Wei, Z., Huang, X., Lu, L., Shangguan, H., Chen, Z., Zhan, J., & Fan, G. (2019). Strategy of rainwater discharge in combined sewage intercepting manhole based onwater quality control. *Water*, *11*(5), 898–911. https://doi.org/10.3390/w11050898

Western Electric Company. (1958). *Statistical quality control handbook* (2nd ed.). Western Electric Co.

Wheeler, D. J. (1983). Detecting of a Shift in Process Average: Tables of the Power Function for X Over Bar Charts. *Journal of Quality Technology*, *15*(4), 155–170. https://doi.org/10.1080/00224065.1983.11978868

Widmer, G., & Kubat, M. (1996). Learning in the presence of concept drift and hidden contexts. *Machine Learning*, *23*(1), 69–101. https://doi.org/10.1023/A:1018046501280

Wikipedia. (n.d.). *Combined sewer*. Retrieved August 20, 2020, from https://en.wikipedia.org/wiki/Combined_sewer

Wojciech, S., Montavon, G., Vedaldi, A., Hansen, L., & Müller, K.-R. (Eds.). (2019). *Explainable AI: Interpreting, Explaining and Visualizing Deep Learning*. Springer.

Xie, Q., Bharat, C., Nazim Khan, R., Best, A., & Hodkiewicz, M. (2017). Cox proportional hazards modelling of blockage risk in vitrified clay wastewater pipes. *Urban Water Journal*, *14*(7), 669–675. https://doi.org/10.1080/1573062X.2016.1236135

Yao, X. (1993). A review of evolutionary artificial neural networks. *International Journal of Intelligent Systems*, *8*(4), 539–567. https://doi.org/10.1002/int.4550080406

Yao, X. (1999). Evolving artificial neural networks. *Proceedings of the IEEE*, *87*(9), 1423–1447. https://doi.org/10.1109/5.784219

Youden, W. J. (1950). Index for rating diagnostic tests. *Cancer*, *3*(1), 32–35. https://doi.org/10.1002/1097-0142(1950)3:1<32::AID-CNCR2820030106>3.0.CO;2-3

Yuan, X., Li, L., Chen, X., & Shi, H. (2015). Effects of precipitation intensity and temperature on ndvi-based grass change over northern china during the period from 1982 to 2011. *Remote Sensing*, *7*(8), 10164–10183. https://doi.org/10.3390/rs70810164

Yue, S., Pilon, P., & Cavadias, G. (2002). Power of the Mann-Kendall and Spearman's rho tests for detecting monotonic trends in hydrological series. *Journal of Hydrology*, *259*, 254–271. https://doi.org/10.1016/S0022-1694(01)00594-7

Zanchetta, A., & Coulibaly, P. (2020). Recent Advances in Real-Time Pluvial

Flash Flood Forecasting. *Water*, *12*(2), 570. https://doi.org/10.3390/w12020570

Zhang, D., Lindholm, G., & Ratnaweera, H. (2018a). DeepCSO: Forecasting of combined sewer overflow at a citywide level using multi-task deep learning. *ArXiv*, *1811.06368*.

Zhang, D., Lindholm, G., & Ratnaweera, H. (2018b). Use long short-term memory to enhance Internet of Things for combined sewer overflow monitoring. *Journal of Hydrology*, *556*, 409–418. https://doi.org/10.1016/j.jhydrol.2017.11.018

Zhang, Q., Li, Z., Snowling, S., Siam, A., & El-Dakhakhni, W. (2019). Predictive models for wastewater flow forecasting based on time series analysis and artificial neural network. *Water Science and Technology*, *80*(2), 243–253. https://doi.org/10.2166/wst.2019.263

Zhao, W., Yu, X., Ma, H., Zhu, Q., Zhang, Y., Qin, W., Ai, N., & Wang, Y. (2015). Analysis of precipitation characteristics during 1957-2012 in the semi-arid loess Plateau, China. *PLoS ONE*, *10*(11), 1–13. https://doi.org/10.1371/journal.pone.0141662

Zweig, M. H., & Campbell, G. (1993). Receiver-operating characteristic (ROC) plots: A fundamental evaluation tool in clinical medicine. *Clinical Chemistry*, *39*(4), 561–577. https://doi.org/10.1093/clinchem/39.4.561

Appendix A: Sensitivity Analyses

This section details the results of several sensitivity analyses conducted to select the modified western electric rules employed in the statistical analysis based and EANN discrepancy-based blockage detection modules, described in Chapter 4.

These analyses were designed to identify a set of run rules which produced a low false positive rate whilst also ensuring a high true positive rate. In particular, a cost-based approach was used to weigh the benefits of true positives against the harm of false positives. In addition, the blockage detection time for the different rules was analysed, to ensure that decreasing the false alarm rate did not excessively increase the detection time of the system.

A.1 Statistical Analyses

A.1.1 Partial Area Under the Curve

The first sensitivity test considered is the partial area under the curve (PAUC) The area under the curve (AUC) is commonly used as a summary measure of the receiver operating characteristic (ROC) curve. A two-class ROC curve is a two-dimensional curve, whereby the true positive rate is plotted on the Y axis and the false positive rate is plotted on the X axis. The ROC curve offers a graphical illustration of the trade-off between the sensitivity and specificity of a test.

A piecewise linear ROC curve is constructed by varying the values of the system and plotting the corresponding points. In general, higher AUC values indicate better test performance. The possible values of the AUC range from 0.5 (no diagnostic ability) to 1 (perfect diagnostic ability).

The partial AUC (PAUC) has been proposed as an alternative measure to the full AUC (Ma et al., 2013). The PAUC considers only the regions of the ROC space where data has been observed or is relevant to the situation. In this case, it is intended that the EDS should produce a low number of false blockage alerts, therefore only rules which produce a low false positive rate are of interest. Thus, only the PAUC with a rate of false positives under 0.02 was analysed.

A.1.2 ROC Curve Cut-off Analysis

Each point on the ROC curve corresponds to a cut-off value and is associated with a test sensitivity and specificity. Although the PAUC is useful for the evaluation of the detection system using different run rules, however it does not specify the optimal cut-off point (i.e. the selected values of the system) directly. Choosing an appropriate cut-off value, which strikes a good compromise between the sensitivity and specificity, is extremely important. There are many approaches for cut-off point selection in the literature (e.g. Liu 2012; Perkins and Schisterman 2006; Unal 2017). The Youden index method (Youden, 1950) is one of the most commonly used techniques. This method defines the optimal cut-point as the point maximizing the Youden function, calculated as the difference between the true positive rate and false positive rate over all possible cut-point values. Another popular approach is the Euclidean index (geometric distance) whereby the cut-off value corresponds to the point on the ROC curve that is closest to the ideal point (0, 1) in the top left hand corner of the ROC space which represents zero false positives and perfect sensitivity (Perkins & Schisterman, 2006).

These approaches are easy to apply and work well, however, they do not take into account the cost of true and false negative results. To overcome this problem a cost-based method has been applied which weighs the benefits of true positives against the harm of false positives. This cost approach is based on an analysis of the costs of the four possible outcomes of a diagnostic test: true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN).

Metz (1978) demonstrated that the slope of the ROC curve at the optimal cut-off value is given as

$$S = \left(\frac{1-P}{P}\right)\left(\frac{C_{FP} - C_{TN}}{C_{FN} - C_{TP}}\right)$$

where P is the positive condition prevalence (i.e. the prevalence of blockage occurrence) and C_{TP}, C_{FN}, C_{FP} and C_{TN} represent the cost for each type.

The point along the ROC curve where the average cost is minimum then corresponds to the cut-off value where

$$f_m = Sensitivity - S(1 - Specificity)$$

is maximized, where f_m is the Cost Index (Zweig & Campbell, 1993).

The Youden index can therefore be seen as a special case of the cost-based approach where the cost ratio is set as one.

A precise determination of the costs and benefits of incorrect and correct classifications is often difficult to obtain – as it is in this situation. However, the exact costs are not needed, instead a qualitative ratio of the costs can be used (Smits, 2010). Following an analysis of data from CSOs in a wastewater network the prevalence of a blockage event was set as 0.01, and based on discussion with industry personal and analysis of the effects of sewer blockage events, the costs were set as $C_{fp} = 2$, $C_{tn} = 0$, $C_{fn} = 1$, $C_{tp} = 0$.

A.2 Results

The sensitivity tests described above were utilised to select the modified western electric rules for the statistical trend based module and the EANN discrepancy based module. The considered western electric rules are presented in Table A-1. The sensitivity tests were used to analyse different combination of these run rules and identify the best performing combination. This process was performed separately to select the rules for the 1, 3 and 6 rainfall category methodologies described in Chapter 4. The selection process is presented here for the 1 rainfall category methodologies only due to space limitations.

The tests were applied to data from 15 different CSO chambers, containing 15 blockage events (both gradual and suddenly forming) and located in both urban and rural catchments. This encompasses a total of 9 years of CSO level and rainfall data. Blockages not detected in under 4 weeks were considered as not detected. This value was agreed with the water utility, bearing in mind that real blockages can remain undetected in sewer systems for many months.

It should be noted that the parameters selected here for the EDS were chosen based on careful analysis of the data and are designed to produce a good performance for all CSO types. However, when the system is deployed by a utility, they will prioritise certain features e.g. a very fast detection time or an extremely low number of false positives. Ultimately, the optimal thresholds will be based on the circumstances in which the system is being employed and on the

operator's desires and expertise. In these cases, the parameters can be easily changed, and the analyses presented here can be used to facilitate an informed decision.

Table A-1 Modified Western Electric Rules for Shewhart control charts.

Run Rule	Description	Constant Multiplier (M)
Rule 1	1 out of 1 consecutive discrepancies fall outside the defined control limits	5
Rule 2	2 out of 3 consecutive discrepancies fall outside the defined control limits	4
Rule 3	4 out of 5 consecutive discrepancies fall outside the defined control limits	3
Rule 4	8 out of 8 consecutive discrepancies fall outside the defined control limits	2
Rule 5	15 out of 15 consecutive discrepancies fall outside the defined control limits	1
Rule 6	25 out of 25 consecutive discrepancies fall outside the defined control limits	0.5

A.2.1 Statistical trend-based analysis module

The selection process for the statistical trend-based analysis module is presented here for the 1 rainfall category methodology. As explained in Chapter 4, The control chart thresholds for each time step t are defined for this module as

If Dry Weather $L_{w,t} = \mu_{dry,w,t} + M_i * N_{dry} * \sigma_{dry,w,t}$

If Wet Weather $L_t = \mu_c + M_i * N_c * \sigma_c$

where μ_c and σ_c are the mean and the standard deviation of the CSO level for rainfall category c of the current timestep. $\mu_{c,w,t}$ and $\sigma_{c,w,t}$ are the mean and the

standard deviation of the CSO level for rainfall category *c* at timestep *t*, with *w* denoting if the timestep is a weekday or weekend. *M* is the constant multiplier defined for each run rule (e.g. 3 in the standard 3-sigma control chart) and N_c is an additional multiplier determined for each rainfall category *c*.

The normalised PAUC, for values with a false positive rate under 0.02, are presented in Table A-2. The PAUC results have been normalised over the area of the curve considered. The ROC curve is presented in Figure A-1. The graph compares different combinations of run rules, with points along each curve representing different values of N from 0.1 to 8. As can been seen both the ROC curve and the normalised PAUC indicate that utilising the higher SPC rules e.g. rules 4, 5 and 6 produce better results.

The Youden index and the Zweig and Campbell index were used to calculate the optimal cut-off points on the ROC curve for each combination of western electric rules. The overall optimal cut-off points according to both methods (i.e. the optimal cut-off point over all the different combinations of rules) are displayed on the ROC curve in Figure A-1. The Youden index indicates that run rules 2-5 produce the best results, indicating an optimum cut off point near to the upper left hand corner of the ROC space, which produces a good true positive rate but an unsatisfactorily high false positive rate. The Zweig and Campbell method, which

Table A-2 Normalised PAUC for different SPC rules for the 1 rainfall category methodology for the statistical trend based analysis module

Run Rule	Normalised PAUC
1	0.45
2	0.47
3	0.52
4	0.57
5	0.66
6	0.71

Run Rule Combination	Normalised PAUC
All Rules	0.68
Rules 1-5	0.61
Rules 2-5	0.54
Rules 2-4	0.62
Rules 3-6	0.69
Rules 4-6	0.70

takes into account the costs of the system, indicates that western electric rules 4 to 6 are superior, selecting a cut-off point closer to the bottom left corner of the ROC space.

Overall, therefore, the tests performed appear to advise selecting WE rules 4-6. However, using this set of rules requires waiting a minimum of 8 to 25 timesteps after evidence of a blockage event is first identified in the CSO level data before raising an alarm - thus significantly increasing the detection time of the system. Requiring this long wait period detrimental to the real-time use of the system by a wastewater utility. Therefore, a compromise was made between minimising the false alarm rate and decreasing the detection time of the system. The chosen rule set and corresponding value of the N was selected as WE rules 3-6 and N=7. The performance of this rule set is displayed in Figure A-1.

The same process was used to select the parameters for the three and six rainfall categories approaches but are not presented here due to space limitations. The selected run rules and corresponding parameters for each of the 3 rainfall category methodologies are detailed in Table A-3. Regarding the selection of the

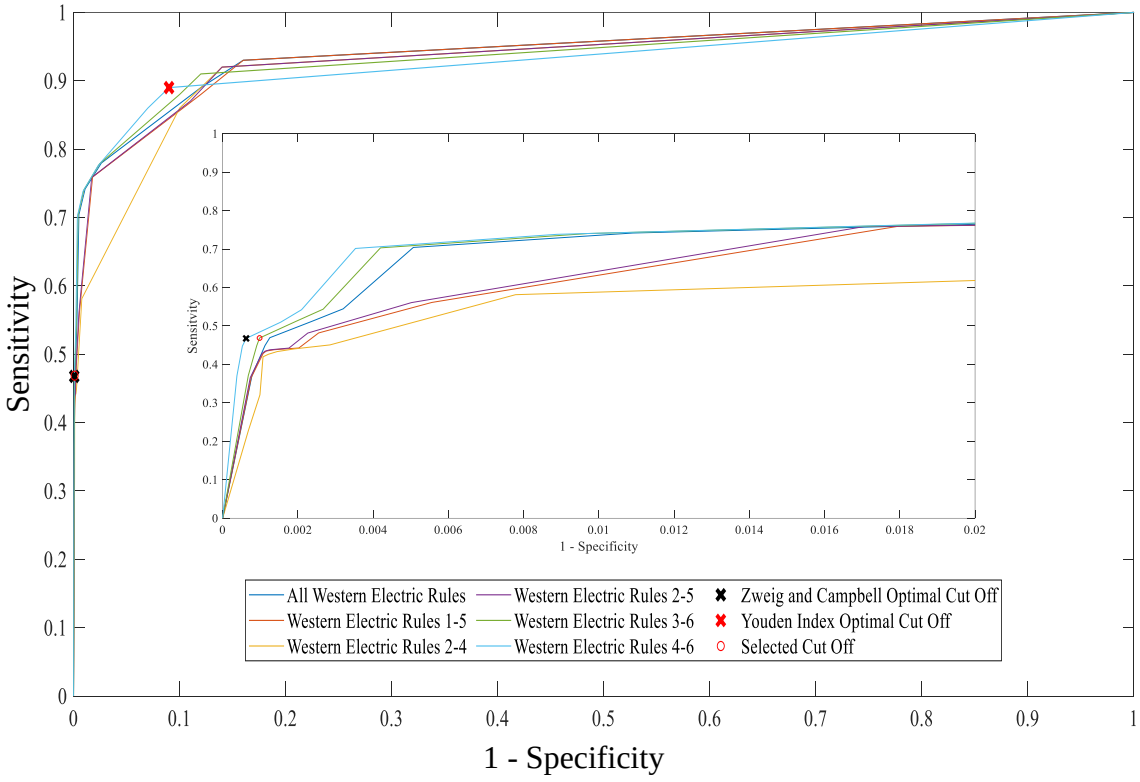


Figure A-1 curves for modified WE run rules for the 1 rainfall category methodology for the statistical analysis based detection module

3 and 6 rainfall category rules, it was found that generally the rainfall classes produced a higher number of false events, despite higher control rule boundaries generated due to the higher values of μ and σ calculated for these classes. Thus, higher values of N and rules which required waiting for additional timesteps before raising an alarm were selected.

Table A-3 Selected modified western electric rules for the statistical analysis based detection module

1 Rainfall Category

Category	Selected SPC rules	N
All Weather	3-6	7

3 Rainfall Categories

Category	Selected SPC rules	N
Dry Weather	2-4	3
Wet Weather	3-6	4
Post Event	2-6	4

6 Rainfall Categories

Category	Selected SPC rules	N
Dry Weather	2-4	3
Rainfall A	2-4	4
Rainfall B	2-4	4
Rainfall C	3-4	5
Rainfall D	3-5	6
Post Event	2-5	4

A.2.2 EANN Discrepancy Based Detection Module

A similar process was applied to the selection of the modified rules for the EANN discrepancy-based model. The selection process for the 1 rainfall category methodology is presented here. The control chart limits for this module are defined as

$$L = \mu_{EANN,c} + M_i * N_c * \sigma_{EANN,c}$$

where $\mu_{EANN,c}$ and $\sigma_{EANN,c}$ are the mean and the standard deviation of the historic EANN discrepancy for rainfall category *c.*

Table A-4 Normalised PAUC for different SPC rules for the 1 rainfall category methodology for the EANN discrepancy based analysis module

Run Rule	Normalised PAUC
1	0.46
2	0.39
3	0.34
4	0.41
5	0.50
6	0.30

Run Rule Combination	Normalised PAUC
All Rules	0.52
Rules 1-5	0.49
Rules 2-5	0.54
Rules 2-4	0.46
Rules 3-6	0.47
Rules 4-6	0.50

Table A-4 displays the normalised PAUC and Figure A-3 presents the ROC for single WE rules and for combinations of run rules. As with the results from the statistical analysis module it can be seen that applying higher rules (4-6) produces superior results in terms of a low false positives rate and high true positive rate. However, these rules also result in a longer blockage detection time. This is demonstrated in Figure A-2 which displays the average detection time vs the false positive rate.

The optimum cut-off points for each combination of run rules according to the Youden index and Zweig and Campbell method were computed, and the overall optimum cut-off points are displayed in Figure A-3 and Figure A-2. As with the parameter selection for the statistical analyses module, the Youden index selects an optimum cut off which produces a good true positive rate but an unsatisfactorily high false positive rate. Therefore, the selected parameters were based on the Zweig and Campbell method, and modified to select a point with a faster detection time. The selected run rules and corresponding parameters are presented in Table A-5. The same process was performed for the selection of the parameters for the 3 and 6 rainfall category methodologies, which are also displayed in Table A-5.

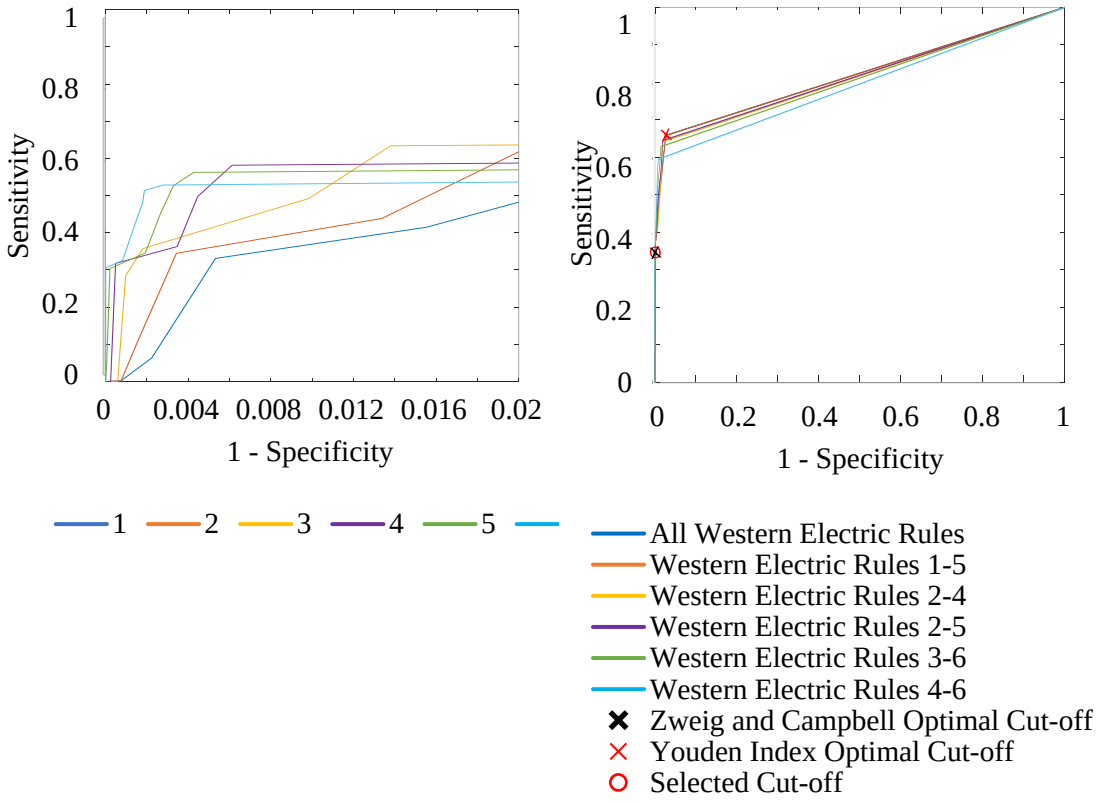


Figure A-3 ROC curves for the 1 rainfall category methodology for the EANN discrepancy based analysis module for (a) for single Western Electric Rules and (b) different combinations of Western Electric rules

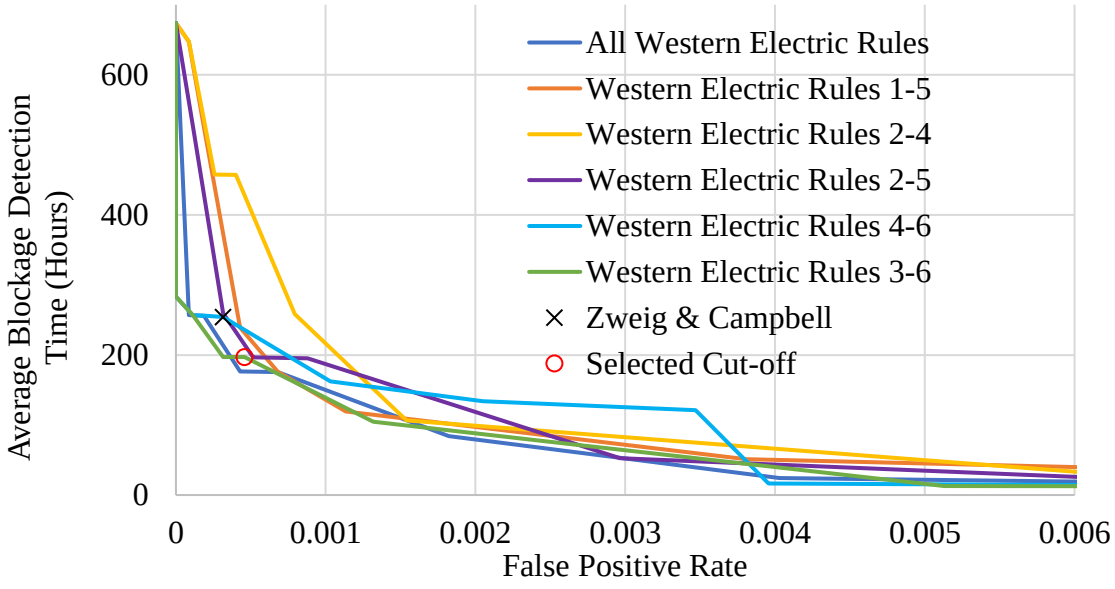


Figure A-2 False Positive Rate vs Average Blockage Detection time for the 1 rainfall category methodology for the EANN discrepancy based analysis module.

Table A-5 Selected modified western electric rules for the EANN discrepancy based detection module

1 Rainfall Category

Category	Selected SPC rules	N
All Weather	3-6	3

3 Rainfall Categories

Category	Selected SPC rules	N
Dry Weather	3-6	3
Wet Weather	4-6	1
Post Event	3-6	1

6 Rainfall Categories

Category	Selected SPC rules	N
Dry Weather	3-6	1
Rainfall A	3-6	0.25
Rainfall B	3-6	0.5
Rainfall C	4-6	2
Rainfall D	4-6	3
Post Event	3-6	2